arXiv:1209.0264v1 [math.NT] 3 Sep 2012

THE TRACE FORMULA AND THE EXISTENCE OF PEL-TYPE ABELIAN VARIETIES MODULO p

ARNO KRET

ABSTRACT. Using the trace formula we show that any Newton stratum of a Shimura variety of PEL-type of type (A) is non-empty at a prime of good reduction.

CONTENTS

INTRODUCTION

Consider a Shimura variety of PEL-type and reduce it modulo a prime p where this variety has good reduction. Then the variety parametrizes Abelian varieties in characteristic p with certain additional PEL-type structures. To each such Abelian variety one may associate its Dieudonné isocrystal. The PEL-structures on the Abelian variety give additional G-structure on the isocrystal, and as such the isocrystals lie in the category of "isocrystals with additional structures" (Kottwitz [15]). We look at these objects modulo isomorphism. When given a Shimura variety of PEL-type modulo p, not every isocrystal with additional G-structure arises from a geometric point on this variety. In fact, there are only a finite number of possible isocrystals; since the work of Rapoport-Richarz and Kottwitz [19, 29] it is known that they all lie in a certain, explicit finite set $B(G_{\mathbb{Q}_p}, \mu)$ of 'admissible' isocrystals, but they did not show that $B(G_{\mathbb{Q}_p}, \mu)$ is *exactly* the set of possible isocrystals, i.e. that for every element $b \in B(G_{\mathbb{Q}_p}, \mu)$ there exists some Abelian variety in characteristic p with additional PEL-type structures whose rational Dieudonné module is equal to b. Recently Wedhorn and Viehmann [36] have proved through geometric means that this is indeed the case if the group

of the Shimura datum is of type (A) or (C). In this article we will show that one can prove this result also using automorphic forms and the trace formula in case the group is of type (A). At the time of writing this article we learned that Sug Woo Shin also found a proof of this theorem with yet another method.

Let us explain our method of proof. The formula of Kottwitz for the number of points on a Shimura variety modulo p can be restricted to count the number of points in any given Newton stratum. Thus, the problem is to show that this formula of Kottwitz does not vanish after restriction. The formula of Kottwitz is in terms of (twisted) orbital integrals on the group G, and he can rewrite the formula in terms of stable orbital integrals on certain endoscopic groups of G. The advantage of this stable expression is that it may be compared to the geometric side of the stable trace formula. By doing so, one will get a certain sum over the endoscopic groups of certain truncated, transferred Hecke operators acting on automorphic representations of these endoscopic groups. (The truncation is defined by the element of $B(G_{\mathbb{Q}_p}, \mu)$.) A general objective is to try and work out this expression; one will then get a rather precise description (of the alternating sum) of the cohomology of the Newton strata. In recent work we have done this for certain simple Shimura varieties, cf. [20], [21]. In this article we have aimed at a simpler goal: We do not try to describe the cohomology of the Newton stratum defined by $b \in B(G_{\mathbb{Q}_p}, \mu)$, we only want to show that this cohomology does not vanish, so that the corresponding Newton stratum must be non-empty. This means that we can pick one very particular Hecke operator f^p and carry out the computation sketched above for this particular Hecke operator. We choose our Hecke operator with care, so that all the proper endoscopy vanishes and that in the end, after applying a simple version of the trace formula, we arrive at a sum of certain b-truncated traces on cuspidal automorphic representations of the quasi-split inner form G^* of the group G (Equation (7.10)):

$$\sum_{\Pi} m(\Pi) \cdot \mathrm{Tr}((f^p)^{G^*}, \Pi^p)\, \mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p). \tag{0.1}$$

In fact, the function f^p will be chosen so that, based on general conjectures, we *expect* that there is precisely *one* automorphic representation Π_0 that contributes to this sum. Therefore no cancellations should occur and the sum is non-zero. We do not try to prove these general conjectures; we show however that, after a sequence of technical lemmas, there is *at least one* contributing representation Π_0, and that for any other hypothetical Π which contributes to Equation (0.1), the quotient

$$\frac{m(\Pi) \cdot \mathrm{Tr}((f^p)^{G^*}, \Pi^p)\, \mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p)}{m(\Pi_0) \cdot \mathrm{Tr}((f^p)^{G^*}, \Pi_0^p)\, \mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_{0,p})}, \tag{0.2}$$

is a positive real number in case α is sufficiently divisible. Then, the sum in Equation (0.1) is non-zero (for α sufficiently divisible). Thus the formula of Kottwitz does not vanish as well, and this means that the corresponding Newton stratum is nonempty. An important step in the argument is to show that Π_0 exists. In particular one has to find a local representation

$\Pi_{0,p}$ at p such that for this local representation we have $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_{0,p}) \neq 0$. In the first section we find a set of such representations with positive Plancherel measure. General theory of automorphic forms will then allow us to find a global automorphic representation Π_0 for which $\Pi_{0,p}$ lies in our Plancherel set.

In upcoming work we will extend the argument to the PEL Shimura varieties of type (C); basically the same argument applies for that case, however some combinatorial computations still have to be completed.

A convention: At various part of the text we will mention integers α which are 'sufficiently divisible' such that a certain statement $(\star)$ holds. With this, we mean to say, by definition, that there exists an integer M such that whenever M divides α, the statement $(\star)$ is true for the integer α.

At some points in the text (when ambiguity is not possible) we will confuse an algebraic group G defined over a field with its set of F-rational points.

Let $x \in \mathbb{R}$ be a real number, then $\lfloor x \rfloor$ (*floor function*) (resp. $\lceil x \rceil$, *ceiling function*) denotes the unique integer in the real interval $(x-1, x]$ (resp. $[x, x+1)$).

Acknowledgements: I would like to thank my thesis advisers Laurent Clozel and Laurent Fargues. Especially Laurent Clozel who helped me putting everything together and encouraged to me that, after the local combinatorial problems have been resolved, the global part of the argument should be "de la cuisine". I thank Erez Lapid for inviting me to come to Jerusalem and explaining me some useful facts on representation theory. And finally I also wish to thank Sug Woo Shin.

1. Isocrystals

We start this preliminary section with some notations. Let p be a prime number and let F be a finite extension of $\mathbb{Q}_p$. Let $\mathcal{O}_F$ be the ring of integers of F, let $\varpi_F \in \mathcal{O}_F$ be a prime element. We write $\mathbb{F}_q$ for the residue field of $\mathcal{O}_F$, and the number q is by definition its cardinality. We fix an algebraic closure $\overline{\mathbb{Q}}_p$ of F, and we let F_α be the unramified extension of F of degree α in $\overline{\mathbb{Q}}_p$. Let G be a smooth reductive group over $\mathcal{O}_F$ (then G_F is an unramified group [33]). We fix a minimal parabolic subgroup P_0 of G, and we standardize the parabolic subgroups of G with respect to P_0. We write $T \subset P_0$ for the Levi component of P_0 and N_0 for the unipotent part, so that we have $P_0 = TN_0$. We call a parabolic subgroup P of G *standard* if it contains P_0, and we write $P = MN$ for its standard Levi decomposition. We write K for the hyperspecial subgroup $G(\mathcal{O}_F) \subset G(F)$. Let $\mathcal{H}(G)$ be the Hecke algebra of locally constant compactly supported complex valued functions on $G(F)$, where the product on this algebra is the one defined by the convolution integral with respect to the Haar measure which gives the group K measure 1. We write $\mathcal{H}_0(G)$ for the spherical Hecke algebra of G with respect to K. We write ρ for the half sum of the positive roots of G.

We write $Z \subset G$ for the center of G, and we write $A \subset Z$ for the split center. Similarly Z_M (resp. Z_P) is the center of the Levi-subgroup M (resp. parabolic subgroup P); and we write A_M (resp. A_P) for the split center of M. We write A_0 for $A_{P_0} \subset T$. We write $\mathfrak{a}_0 := X_*(A_0) \otimes \mathbb{R}$, and C_0 for the closed, positive chamber in $\mathfrak{a}_0$:

$$C_0 := \{x \in \mathfrak{a}_0 \mid \text{for all roots } \alpha \text{ in } \Delta(A_0, \mathrm{Lie}(N_0))\colon \langle x, \alpha \rangle \geq 0\}.$$

Let $B(G)$ be the set of σ-conjugacy classes in $G(L)$, where L is the completion of the maximal unramified extension of F and σ is the arithmetic Frobenius of L over F. Let $\mu \in X_*(T)$ be a G-dominant minuscule cocharacter. Recall that Kottwitz has defined the subset $B(G, \mu) \subset B(G)$ of μ-admissible isocrystals [19, 29].

Let $\mathbb{D}$ be the protorus over F with character group given by $X_*(\mathbb{D}) = \mathbb{Q}$ and trivial Galois action. For any $b \in G(L)$ we have an unique morphism $\nu_b\colon \mathbb{D}_L \to G_L$ characterized by the following property: For every algebraic representation (ρ, V) of G on a finite dimensional vector space V the composition $\rho \circ \nu_b$ determines the slope filtration on $(V \otimes L, \rho(b)(1 \otimes \sigma_L))$ [15, §4]. Replacing b by a σ-conjugate amounts to conjugating ν_b with some $G(L)$-conjugate. Moreover, one can replace b so that ν_b has image inside the torus $A_{0,L}$, so that ν_b defines an element of $\mathfrak{a}_0$ [19, p. 267] [29, 1.7]. Write $\overline{\nu}_b$ for the unique element of C_0 whose orbit under the Weyl group meets ν_b. The morphism $\overline{\nu}_b$ is called the *slope morphism* and the mapping $B(G) \to C_0, b \mapsto \overline{\nu}_b$ is called the *Newton map*. Note that the mapping $b \mapsto \overline{\nu}_b$ is not injective in general (it is injective in case $G = \mathrm{GL}_n(F)$).

Recall that we fixed an embedding $F \subset \overline{\mathbb{Q}}_p$. For each finite subextension $F' \subset \overline{\mathbb{Q}}_p$ of F we have the unique mapping $H_T\colon T(F') \to X_*(T)_{\mathbb{R}}$ such that $q_{F'}^{-\langle \chi, H(t) \rangle} = |\chi(t)|$ for all $t \in T(F')$, where $q_{F'}$ is the cardinal of the residue field of F', and the norm is normalized so that $|p|$ equals $q_{F'}^{-e}$ where e is the ramification index of F'/F. By taking the union over all F' we get a mapping $H_T\colon T(\overline{\mathbb{Q}}_p) \to X_*(T)_{\mathbb{R}}$. Consider the composition H_A defined by $T(\overline{\mathbb{Q}}_p) \to X_*(T)_{\mathbb{R}} \to X_*(A)_{\mathbb{R}} = \mathfrak{a}_0$. Let $G(\overline{\mathbb{Q}}_p)_{\mathrm{ss}} \subset G(\overline{\mathbb{Q}}_p)$ be the subset of semisimple elements. If $g \in G(\overline{\mathbb{Q}}_p)_{\mathrm{ss}}$, then we may conjugate g to an element g' of $T(\overline{\mathbb{Q}}_p)$ and then consider $H_A(g') \in \mathfrak{a}_0$. This element of $\mathfrak{a}_0$ is only defined up to conjugacy, but we can take a representative in the, closed positive Weyl chamber $H(g) \in C_0^+$ which is well-defined. Thus we have a map $\Phi\colon G(\overline{\mathbb{Q}}_p)_{\mathrm{ss}} \to C_0$ defined on the semisimple elements. We extend the definition of Φ to $G(\overline{\mathbb{Q}}_p)$ by defining $\Phi(g) := \Phi(g_{\mathrm{ss}})$, where g_{ss} is the semisimple part of the element $g \in G(\overline{\mathbb{Q}}_p)$. We can restrict to $G(F) \subset G(\overline{\mathbb{Q}}_p)$ to obtain the mapping $\Phi\colon G \to C_0$. In Proposition 1.1 we establish a relation between the map Φ and the Newton polygon mapping of isocrystals.

We recall the definition of the norm $\mathcal{N}$ of (certain) σ-conjugacy classes (cf. [3] [13, p. 799]). To any element $\delta \in G(F_\alpha)$ we associate the element $N(\delta) := \delta\sigma(\delta)\cdots\sigma^{\alpha-1}(\delta) \in G(F_\alpha)$. For any element $\delta \in G(F_\alpha)$, defined up to σ-conjugacy, with semi-simple norm $N(\delta)$ one proves (see [*loc. cit.*]) that $N(\delta)$ actually comes from a conjugacy class $\mathcal{N}(\delta)$ in the group $G(F)$.

Proposition 1.1. *Let α be a positive integer and let $\delta \in G(F_\alpha)$ be an element of semi-simple norm, defined up to σ-conjugacy. Let $\gamma \in G(F)$ be an element in the conjugacy class $\mathcal{N}(\delta)$, and let b be the isocrystal with additional G-structure defined by δ. Then $\overline{\nu}_b = \alpha \cdot \Phi(\gamma) \in C_0$.*

Proof. We first prove the case where G is the general linear group. If $G = \mathrm{GL}_{n,F}$, then an isocrystal "with additional G-structure" is simply an isocrystal, i.e. a pair (V, Φ) where V is an n-dimensional L vector space and Φ is a σ-linear bijection from V onto V. Because b is induced by some $\delta \in G(F_\alpha)$, we may find a F_α-vector space V' together with a σ-linear bijection $\Phi' \colon V' \to V'$ such that (V, Φ) is obtained from (V, Φ) by extending the scalars $V = V' \otimes_{F_\alpha} L$ and $\Phi(v' \otimes l) := \Phi'(v') \otimes \sigma(l)$. Then (V', Φ') is an F_α-space in the terminology of Demazure [11], and a theorem of Manin gives the relation $\overline{\nu}_b = \alpha \cdot \Phi(\gamma)$ (cf. [11, p. 90]).

Drop the assumption that $G = \mathrm{GL}_n$. Pick a faithful representation $\rho \colon G \to \mathrm{GL}_V$ of G in some finite dimensional $\mathbb{Q}_p$-vector space V. Then, by the statement for GL_n, we see that $\alpha \cdot \Phi_{\mathrm{GL}_n}(\rho(\gamma))$ determines the slope filtration on the space $(V \otimes L, \rho(b)(1 \otimes \sigma_L))$. Thus $\rho \circ \overline{\nu}_b = \alpha \cdot \Phi_{\mathrm{GL}_n}(\rho(\gamma))$, and then the equality is also true for the group G. □

We now study the set $B(G)$, where G is an unramified unitary group over F, which therefore splits over the extension F_2/F. The absolute root system of G is isomorphic to the usual root system in $\mathbb{R}^n$ of type A (cf. Bourbaki [4, chap. 6]), and the non-trivial element of the group $\mathrm{Gal}(F_2/F)$ acts on $\mathbb{R}^n$ via the operator θ defined by $(x_1, x_2, \ldots, x_n) \longmapsto (-x_n, -x_{n-1}, \ldots, -x_1)$. The space $\mathfrak{a}_0$ is the subspace of θ invariant elements in $\mathbb{R}^n$, thus it is equal to the set of $(x_i) \in \mathbb{R}^n$ with $x_i = -x_{n+1-i}$ for all indices i. The dimension of this space is equal to $\lfloor n/2 \rfloor$.

Whenever $b \in B(G)$ is an isocrystal with G-structure, we have its slope morphism $\overline{\nu}_b \in C_0$, which we may view as an θ-invariant element of $\mathbb{R}^n$. This way we get the slopes $\lambda_1, \lambda_2, \ldots, \lambda_n$ of b, which are just the coordinates of the vector $\overline{\nu}_b \in \mathbb{R}^n$. We order them so that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. These slopes satisfy the property $\lambda_i = -\lambda_{n+1-i}$. We associate to these slopes the Newton polygon $\mathcal{G}_b$ of b, which is by definition the continuous piecewise linear function from the real interval $[0, n]$ to $\mathbb{R}$ with the property that the only points where it is possibly not differentiable are the integral points $[0, n] \cap \mathbb{Z}$; the value of $\mathcal{G}_b$ at these points is defined by: $\mathcal{G}_b(0) := 0$ and $\mathcal{G}_b(i) := \lambda_1 + \lambda_2 + \ldots + \lambda_i$. Observe that, due to the θ-invariance, we have $\mathcal{G}_b(n) = \lambda_1 + \ldots + \lambda_n = 0$. Furthermore the graph (or polygon) $\mathcal{G}_b$ is symmetric around the vertical line that goes through the point $(\frac{n}{2}, 0)$. In Figure 1 we show a typical unitary Newton polygon. In particular negative slopes may occur, which does not happen for the group $\mathrm{GL}_n(F)$ nor for the group $\mathrm{Gsp}_{2g}(F)$.

Let us now determine what the Hodge polygons looks like. The minuscule cocharacter μ is defined over $\overline{F}$, and is given by

$$\mu = (\underbrace{0, 0, \ldots, 0}_{n-s}, \underbrace{1, 1, \ldots, 1}_{s}) \in \mathbb{Z}^n \subset \mathbb{R}^n,$$

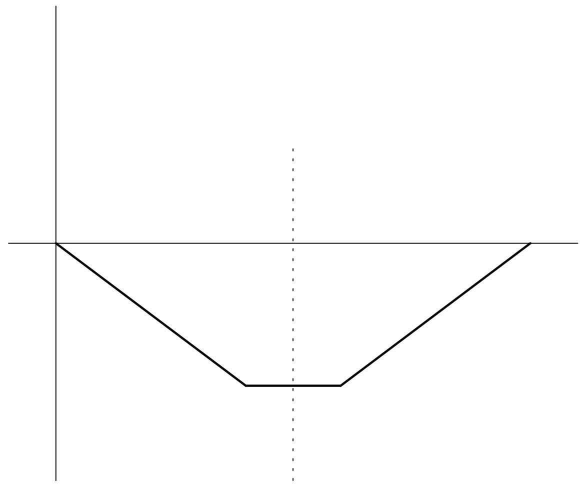

FIGURE 1. The dark line is an example of the Newton polygon of an isocrystal b with additional U_{10}^*-structure. The horizontal line from $(0,0)$ to $(10,0)$ is the Newton polygon of the basic isocrystal. The vertical dotted line indicates the mirror symmetry of the Newton polygons of the G-isocrystals.

for some integer s with $0 \leq s \leq n$. To define the set $B(G,\mu)$ Kottwitz [19, §6] takes the average of μ under the Galois action to get

$$\overline{\mu} := \tfrac{1}{2}(\mu + \theta(\mu)) = (\underbrace{-\tfrac{1}{2}, -\tfrac{1}{2}, \ldots, -\tfrac{1}{2}}_{s'}, \underbrace{0, 0, \ldots, 0}_{n-2s'}, \underbrace{\tfrac{1}{2}, \tfrac{1}{2}, \ldots, \tfrac{1}{2}}_{s'}) \in \mathfrak{a}_0 \subset \mathbb{R}^n,$$

where $s' := \min(s, n-s)$. To this element $\overline{\mu} \in \mathbb{R}^n$ we may associate in the same manner a graph $\mathcal{G}_\mu$ as in Figure 1. Then $b \in B(G)$ lies in $B(G,\mu)$ if and only if the end point of $\mathcal{G}_b$ is $(n,0)$ and if $\mathcal{G}_b$ lies above[1] the graph $\mathcal{G}_\mu$.

2. PEL DATUM

Let $G/\mathbb{Q}$ be a unitary group of similitudes which arises from a PEL type Shimura datum [18, §5]. We recall briefly the definition of G from [*loc. cit.*]. Let $B/\mathbb{Q}$ be a finite dimensional simple algebra and write F for its center, and assume that F is a CM field. Let $*$ be a positive involution on B over $\mathbb{Q}$ which induces on F the complex conjugation. Write $F^+ \subset F$ for the fixed field of $*$ on F. Let V be a nonzero finitely generated left B-module. Let $(\cdot,\cdot)$ be a nondegenerate $\mathbb{Q}$-valued alternating form on V such that $(bv, w) = (v, b^*w)$ for all $v, w \in V$ and all $b \in B$. Then $G/\mathbb{Q}$ is the algebraic group with for all commutative $\mathbb{Q}$-algebras R:

$$G(R) = \left\{ g \in \mathrm{End}_B(V)^\times | \exists c(g) \in R^\times : (g\cdot, g\cdot) = c(g)(\cdot,\cdot) \text{ on V} \right\}. \tag{2.1}$$

Let $G_1 \subset G$ be the kernel of the similitudes ratio. Then G_1 is obtained by restriction of scalars of a unitary group G_0 defined over the totally real field F^+ (following the notations

[1] Lies above in the non-strict sense, the two graphs may touch, or even be the same (the ordinary case).

of [*loc. cit.*]. The group $G_{1,\mathbb{Q}_p}$ is isomorphic to a product of groups

$$G_{1,\mathbb{Q}_p} \cong \prod_{\wp | p} G_{1,\wp}, \tag{2.2}$$

where $\wp$ ranges over the F^+-places above p, and where the group $G_{1,\wp}$ can be either the restriction of scalars to $\mathbb{Q}_p$ of $\mathrm{GL}_{n,F^+_\wp}$ or of an unramified unitary group over $F^+_\wp$. We will study the group $G_{1,\mathbb{Q}_p}$ factor by factor. Thus, in this paper we will need to work not only with unramified unitary groups, but with the slightly more general class of groups of the form $\mathrm{Res}_{F'_\wp/F_\wp} U$, where $F'_\wp/F_\wp$ is some unramified extension and U is an unramified unitary group over $F'_\wp$. The study of isocrystals over these groups reduces quickly to the study of isocrystals over the group U (which we did above), by the Shapiro bijection (cf. [19, 6.5.3]):

$$B(\mathrm{Res}_{F'_\wp/F_\wp} U) = B_{F'_\wp}(U), \tag{2.3}$$

where we have added the subscript "$F'_\wp$" in the right hand side to indicate that there we work with σ'-conjugacy classes, where σ' is the arithmetic Frobenius of $\overline{\mathbb{Q}}_p$ over $F'_\wp$. Under the Shapiro bijection the subset $B(\mathrm{Res}_{F'_\wp/F_\wp} U, \mu_\wp)$ corresponds to the subset $B_{F'_\wp}(U, \mu'_\wp)$ of $B_{F'_\wp}(U)$, where $\mu'_\wp$ is defined by

$$\mu'_\wp \stackrel{\mathrm{def}}{=} \sum_{v \in V(\wp)} (\underbrace{1, 1, \ldots, 1}_{s_v}, \underbrace{0, 0, \ldots, 0}_{n - s_v}) \in \mathbb{Z}^n.$$

Thus, the combinatorics for isocrystals with $\mathrm{Res}_{F'/F} U$-structure is almost the same as the combinatorics for the case $F' = F$; only the Hodge polygons are slightly more complicated.

We recall briefly how the functions of Kottwitz ϕ_α and f_α are constructed [18, §5] [16, p. 173]. Let E be the reflex field, let p be a prime number where the Shimura variety has good reduction in the sense of [18, §6]. In particular the field E is unramified at p; let $\mathfrak{p}$ be an E-prime above p. Write $E_\mathfrak{p}$ for the completion of E at $\mathfrak{p}$, fix an embedding $E_\mathfrak{p} \subset \overline{\mathbb{Q}}_p$ and let for each positive integer α, the field $E_{\mathfrak{p},\alpha} \subset \overline{\mathbb{Q}}_p$ be the unramified extension of $E_\mathfrak{p}$ of degree α. In the PEL datum there is fixed a $*$-morphism $h\colon \mathbb{C} \to \mathrm{End}(B)^{\mathrm{opp}}_\mathbb{R}$. This morphism induces a morphism of algebraic groups from Deligne's torus $\mathrm{Res}_{\mathbb{C}/\mathbb{R}} \mathbb{G}_\mathrm{m}$ to the group $G_\mathbb{R}$. Tensor this morphism with $\mathbb{C}$ to get a morphism from $\mathbb{G}_\mathrm{m} \times \mathbb{G}_\mathrm{m}$ to $G_\mathbb{C}$ and then restrict to the factor $\mathbb{G}_\mathrm{m}$ of the product $\mathbb{G}_\mathrm{m} \times \mathbb{G}_\mathrm{m}$ corresponding to the identity $\mathbb{R}$-isomorphism $\mathbb{C} \to \mathbb{C}$. This way we obtain a cocharacter $\mu \in X_*(G)$. We quote from Kottwitz's paper at Ann Arbor, p. 173: The $G(\mathbb{C})$ conjugacy class of μ gives a $G(\overline{\mathbb{Q}}_p)$ conjugacy class of morphisms which are fixed by the Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}_p/E_{\mathfrak{p},\alpha})$. Let S_α be a maximal $E_{\mathfrak{p},\alpha}$-split torus in the group G over the ring of integers $\mathcal{O}_{E_{\mathfrak{p},\alpha}}$. By Lemma (1.1.3) of [14] we can choose μ so that it factors through S_α. Then $\phi_\alpha = \phi_{G,\mu,\alpha}$ is the characteristic function of the double coset $G(\mathcal{O}_{\mathfrak{p},\alpha})\mu(p^{-1})G(\mathcal{O}_{\mathfrak{p},\alpha})$. The function $f_\alpha = f_{G,\mu,\alpha}$ is by definition the base change [3, 26] of ϕ_α from the group $G(E_{\mathfrak{p},\alpha})$ to the group $G(\mathbb{Q}_p)$.

3. Truncated traces

We revert to the general notations of the beginning of the first section, thus G is a connected unramified reductive group over a local field. In this section we introduce the concept of truncated traces of smooth representations with respect to elements of the set $B(G)$, i.e. the isocrystals with additional G-structure. We will then compute these truncated traces on the Steinberg representation and on the trivial representation.

Using the mapping Φ from the previous section we define the truncated traces with respect to an arbitrary element $b \in B(G)$:

Definition 3.1. Let $\nu \in C_0$. We define:

$$\Omega_\nu^G \stackrel{\text{def}}{=} \{g \in G \mid \exists \lambda \in \mathbb{R}_{>0} : \ \Phi(g) = \lambda \cdot \nu \in C_0\}. \tag{3.1}$$

We let χ_ν^G be the characteristic function on of the subset Ω_ν^G of G. Let $b \in B(G)$ be an isocrystal with additional G-structure. Then we will write $\chi_b^G := \chi_{\overline{\nu}_b}^G$ and $\Omega_b^G := \Omega_{\overline{\nu}_b}^G$.

Remark. The Newton mapping $B(G) \ni b \mapsto \overline{\nu}_b \in C_0$ is *injective* for a simply connected, connected quasi-split reductive group over a non-Archimedean local field [19, §6].

Let $P = MN$ be a standard parabolic subgroup of G and let A_P be the split center of P, we write $\varepsilon_P = (-1)^{\dim(A_P/A_G)}$. To the parabolic subgroup P we associate the subset $\Delta_P \subset \Delta$ consisting of those roots which act non trivially on A_P. Define $\mathfrak{a}_P$ to be $X_*(A_P)_\mathbb{R}$ and define $\mathfrak{a}_P^G$ to be the quotient of $\mathfrak{a}_P$ by $\mathfrak{a}_G$, and define $\mathfrak{a}_P^+$ by

$$\mathfrak{a}_P^+ := \{x \in \mathfrak{a}_P \mid \text{for all roots } \alpha \text{ in } \Delta_P\text{: } \langle x, \alpha\rangle > 0\}.$$

We recall the definition of the obtuse and acute Weyl-chambers [25, 34]. Let P be a standard parabolic subgroup of G. We write $\mathfrak{a}_0 = \mathfrak{a}_{P_0}$ and $\mathfrak{a}_0^G = \mathfrak{a}_{P_0}^G$. For each root α in Δ we have a coroot $\alpha^\vee$ in $\mathfrak{a}_0^G$. For $\alpha \in \Delta_P \subset \Delta$ we send the coroot $\alpha^\vee \in \mathfrak{a}_0^G$ to the space $\mathfrak{a}_P^G$ via the canonical surjection $\mathfrak{a}_0^G \twoheadrightarrow \mathfrak{a}_P^G$. The set of these restricted coroots $\alpha^\vee|_{\mathfrak{a}_P^G}$ with α ranging over Δ_P form a basis of the vector space $\mathfrak{a}_P^G$. By definition the set of fundamental roots $\{\varpi_\alpha \in \mathfrak{a}_P^{G*} \mid \alpha \in \Delta_P\}$ is the basis of $\mathfrak{a}_P^{G*} = \mathrm{Hom}(\mathfrak{a}_P^G, \mathbb{R})$ dual to the basis $\{\alpha^\vee_{\mathfrak{a}_P^G}\}$ of coroots. We let τ_P^G be the characteristic function on the space $\mathfrak{a}_P^G$ of the *acute Weyl chamber*,

$$\mathfrak{a}_P^{G+} = \left\{x \in \mathfrak{a}_P^G \mid \forall \alpha \in \Delta_P \ \langle \alpha, x\rangle > 0\right\}. \tag{3.2}$$

We let $\widehat{\tau}_P^G$ be the characteristic function on $\mathfrak{a}_P^G$ of the *obtuse Weyl chamber*,

$${}^+\mathfrak{a}_P^G = \left\{x \in \mathfrak{a}_P^G \mid \forall \alpha \in \Delta_P \ \langle \varpi_\alpha^G, x\rangle > 0\right\}. \tag{3.3}$$

We define the function χ_N to be the composition $\tau_P^G \circ (\mathfrak{a}_P \twoheadrightarrow \mathfrak{a}_P^G) \circ H_M$, and we define the function $\widehat{\chi}_N$ to be the composition $\widehat{\tau}_P^G \circ (\mathfrak{a}_P \twoheadrightarrow \mathfrak{a}_P^G) \circ H_M$. The functions χ_N and $\widehat{\chi}_N$ are locally constant and K_M-invariant, where $K_M = M(\mathcal{O}_F)$.

Let $b \in B(G)$ be an isocrystal with additional G structure and let $\overline{\nu}_b \in C_0$ be its slope morphism. For any standard parabolic subgroup $P \subset G$ we have the subset $\mathfrak{a}_P^+ \subset C_0$. Let

P_b be the standard parabolic subgroup of G such that $\overline{\nu}_b \in \mathfrak{a}_{P_b}^+$. We call the group P_b the subgroup of G *contracted* by the isocrystal $b \in B(G)$. These groups are precisely the parabolic subgroups which appear in the Kottwitz decomposition of the set $B(G)$ (see [19, 5.1.1]). We write $P_b = M_b N_b$ for the standard decomposition of P_b.

We write π_{0P} for the projection from the space $\mathfrak{a}_0$ onto $\mathfrak{a}_P$, it sends an element $X \in \mathfrak{a}_0$ to its average under the action of the Weyl group.

We introduce a certain characteristic function on G associated to the isocrystal $b \in B(G)$:

Definition 3.2. Let $P_b = M_b N_b$ be the standard parabolic subgroup of G contracted by b. We define η_b to be the characteristic function on G of the set of elements $g \in G$ such that there exists a $\lambda \in \mathbb{R}_{>0}^\times$ such that $\pi_{0P}(\Phi(g)) = \lambda \cdot \overline{\nu}_b \in \mathfrak{a}_P^+$.

Remark. If the isocrystal b is basic, then we have $P = G$, and the element $\overline{\nu}_b \in C_0$ is central. Therefore the function η_b is *spherical*.

In case the isocrystal $b \in b(G)$ is basic then χ_b^G coincides with $\eta_b \chi_c^G$:

Lemma 3.3. *Let $b \in B(G)$ be a basic isocrystal. Then we have $\chi_b^G = \eta_b \chi_c^G$.*

Proof. Let $g \in G$, and consider $\Phi(g) \in C_0$. Then g is compact if and only if it contracts G as parabolic subgroup which means that $\Phi(g)$ lies in $\mathfrak{a}_G \subset C_0^+$. Assume g is compact. Then $\chi_b^G(g) = 1$ if and only if the slope morphism $\overline{\nu}_b$ of b lies in $\mathfrak{a}_G$, i.e. if and only if the centralizer of the slope morphism of b is equal to G. But that means that b is basic. Conversely, assume b is basic. Then its slope morphism is central, thus $\chi_c^G(g) = 1$ if and only if g contracts G, i.e. g is compact. Furthermore we have $\eta_b(g) = 1$ because $\Phi(g)$ equals $\overline{\nu}_b$ up to a positive scalar. This completes the proof. □

We call the collection of subsets Ω_b^G for $b \in B(G)$ the *Newton polygon stratification* of the group G. For our proofs we will also need to study another stratification, called the *Casselman stratification* of G:

Definition 3.4. Let Q be a standard parabolic subgroup of G. We define $\Omega_Q^G \subset G$ to be the subset of elements $g \in G$ which contract [6, §1] a parabolic subgroup conjugate to Q. Write χ_Q^G for the characteristic function on G of the subset $\Omega_Q^G \subset G$. These sets Ω_Q^G form the Casselman stratification of G.

For truncated traces with respect to the Casselman stratification we have:

Proposition 3.5. *Let $Q = LU$ be a standard parabolic subgroup of G. Let $f \in \mathcal{H}(G)$ be a locally constant function with compact support. Then we have* $\mathrm{Tr}(\chi_Q^G f, \pi) = \mathrm{Tr}(\chi_U \chi_c^L \overline{f}^{(Q)}, \pi_U(\delta_Q^{-1/2}))$.

Proof. By the Proposition [6, prop 1.1] on compact traces, for all functions f on G, the full trace $\mathrm{Tr}(f, \pi)$ is equal to the sum of compact traces $\sum \mathrm{Tr}_M(\chi_c^M \overline{f}^{(P)}, \pi_N(\delta_P^{-1/2}))$, where the

sum ranges over the standard parabolic subgroups $P = MN$ of G. Consider only those functions of the form $\chi_Q^G f \in \mathcal{H}(G)$. Then we obtain that the trace $\mathrm{Tr}(\chi_Q^G f, \pi)$ is equal to the sum $\sum \mathrm{Tr}_M(\chi_c^M \chi_Q^G \overline{f}^{(P)}, \pi_N(\delta_P^{-1/2}))$ in which $P = MN$ ranges over the standard parabolic subgroups of G. Observe that $\chi_c^M \chi_Q^G = 0$ if $P \neq Q$. Therefore only the term corresponding to $P = Q$ remains in the sum. This completes the proof. $\square$

Let us now explain the relation between the Casselman stratification and the Newton stratification. The following Proposition gives the relation between the Casselman stratification of G and the Newton stratification:

Proposition 3.6. *For all $b \in B(G)$ we have $\Omega_b^G \subset \Omega_{P_b}^G$.*

Proof. Assume that $g \in \Omega_b^G$. Then $\Phi(g) = \lambda \overline{\nu}_b \in \mathfrak{a}_0$. Let P be the standard parabolic subgroup of G which is conjugate to the parabolic subgroup of G contracted by g. Then $\overline{\nu}_b = \lambda \Phi(g) \in \mathfrak{a}_P^+$. Then, by definition, P is the parabolic subgroup contracted by b. This completes the proof. $\square$

Example. The inclusion $\Omega_b^G \subset \Omega_{P_b}^G$ is strict in general. Consider for example the case $G = \mathrm{GL}_{n,\mathbb{Q}_p}$ to see that it is non-strict only in particular cases, such as when $n = 2$. In the particular case of the Shimura varieties of Harris-Taylor [12] one can also use the Casselman stratification to separate the isocrystals.

We will now compute the truncated trace on the Steinberg representation.

Definition 3.7. Let ξ_b^{St} be the characteristic function on T defined by $\xi_b^{St} := \widehat{\chi}_{N_0 \cap M_b} \chi_{N_b} \eta_b$, where with the notation $\widehat{\chi}_{N_0 \cap M_b}$ we mean the characteristic function on the Levi subgroup $M_b \subset G$, corresponding to the obtuse chamber relative to the minimal parabolic subgroup of M_b.

Proposition 3.8. *Let $f \in \mathcal{H}_0(G)$ be a spherical Hecke operator. Then we have*

$$\mathrm{Tr}(\chi_b^G f, \mathrm{St}_G) = \varepsilon_{P_0 \cap M_b} \mathrm{Tr}_T(\xi_b^{St} f^{(P_0)}, \mathbf{1}(\delta_{P_b}^{-1/2} \delta_{P_0 \cap M_b}^{1/2})). \tag{3.4}$$

Proof. Write $P = MN$ for the parabolic subgroup contracted by the isocrystal b. We compute:

$$\mathrm{Tr}(\chi_b^G f, \mathrm{St}_G) = \mathrm{Tr}_M(\chi_b^G \chi_N f^{(P)}, (\mathrm{St}_G)_N(\delta_P^{-1/2})), \tag{3.5}$$

(Proposition 3.5). Let $b_M \in B(M)$ be a G-regular basic element such that its image in $B(G)$ is equal to b [15, prop. 6.3]. By [*loc. cit.*] the set of all such b_M are G-conjugate. As functions on M we have $\chi_b^G \chi_N = \chi_{b_M}^M \chi_N$. Therefore we may simplify Equation (3.5) to $\mathrm{Tr}_M(\chi_{b_M}^M \chi_N f^{(P)}, (\mathrm{St}_G)_N(\delta_P^{-1/2}))$. By Lemma 3.3 the latter trace equals $\mathrm{Tr}_M(\chi_c^M \eta_b \chi_N f^{(P)}, (\mathrm{St}_G)_N(\delta_P^{-1/2}))$. In our article [20] we computed the compact traces on the Steinberg representation for all spherical Hecke operators. By [20, Prop. 1.13] we get

$$\mathrm{Tr}(\chi_b^G f, \mathrm{St}_G) = \varepsilon_{P_0 \cap M} \mathrm{Tr}(\widehat{\chi}_{N_0 \cap M} \eta_b \chi_N f^{(P_0)}, \mathbf{1}(\delta_P^{-1/2} \delta_{P_0 \cap M}^{1/2})). \tag{3.6}$$

This completes the proof. $\square$

In the same way one may compute the truncated traces on the trivial representation. We have to introduce two more notations. Let $\widehat{\chi}^{\leq}_{N_0 \cap M_b}$ be the characteristic function on M_b corresponding to the negative closed obtuse chamber in $\mathfrak{a}_P$. Then we define:

Definition 3.9. Let $b \in B(G)$ be an isocrystal. We define $\xi_b^{\mathbf{1}} := \widehat{\chi}^{\leq}_{N_0 \cap M_b} \chi_{N_b} \eta_b$.

Proposition 3.10. *We have* $\mathrm{Tr}(\chi_b^G f, \mathbf{1}) = \mathrm{Tr}_T(\xi_b^{\mathbf{1}} f^{(P_0)}, \mathbf{1}(\delta_{P_0}^{-1/2}))$.

Proof. The proof of Proposition 3.8 may be repeated without change up to Equation (3.6). Replace the result in that last Equation with the result from Proposition 3.1 from [21] which computes the compact trace on the trivial representation for any Hecke operator (and any unramified group). □

Remark. With a method similar to the above one may compute the truncated traces on the irreducible subquotients of the G-representation on the space $C^\infty(G/P_0)$ of locally constant functions on G/P_0.

4. The class of $\mathfrak{R}(b)$-representations

For the global applications to Shimura varieties we find a class representations $\mathfrak{R}_1(b)$ of positive Plancherel density on which the truncated trace of the Kottwitz functions are non-zero. In fact we will show that for *most* of the isocrystals $b \in B(G, \mu)$ one can simply take the Steinberg representation at p, but there are some exceptions where the truncated trace on the Steinberg representation vanishes, in those cases we have to take a different representation.

Let G be a connected, reductive unramified group over $\mathbb{Q}_p$, let P_0 be a Borel subgroup of G. Let T be the Levi-component of P_0. Then T is a maximal torus in G, and let W be the absolute Weyl group of T in G. Let $\mu \in X_*(T)$ be a minuscule cocharacter.

We write in this section E for an arbitrary, finite unramified extension of $\mathbb{Q}_p$. In later sections, the field E that we consider here will be the completion of the reflex field at a prime of good reduction. We fix an embedding of E into $\overline{\mathbb{Q}}_p$, and for each positive integer α we write $E_\alpha \subset \overline{\mathbb{Q}}_p$ for the unramified extension of degree α of E.

Definition 4.1. (cf. [14]). Let α be a positive integer, and E_α the unramified extension of E of degree α contained in $\overline{\mathbb{Q}}_p$. We write W_α for the subgroup $W(G(E_\alpha), T(E_\alpha))$ of W. Write S_α for a maximal E_α-split subtorus of G_{E_α}. We define $\phi_{G,\mu,\alpha} \in \mathcal{H}_0(G(E_\alpha))$ to be the spherical function whose Satake transform is equal to

$$p^{-\alpha\langle \rho_G, \mu\rangle} \sum_{w \in W_\alpha/\mathrm{stab}_{W_\alpha}(\mu)} [w(\mu)] \in \mathbb{C}[X_*(S_\alpha)]^{W_\alpha}, \tag{4.1}$$

where $\mathrm{stab}_{W_\alpha}(\mu) \subset W_\alpha$ is the stabilizer of μ in the group W_α. We define $f_{G,\mu,\alpha}$ to be the function which is obtained from $\phi_{G,\mu,\alpha}$ via base change from the group $G(E_\alpha)$ to the group $G(F^+)$. We call $f_{G,\mu,\alpha}$ the *function of Kottwitz*.

Remark. Kottwitz proves in [14] that the definition of the Kottwitz functions $f_{G,\mu,\alpha}$ and $\phi_{\mathbb{G},\mu,\alpha}$ coincide with the definition that we gave at the end of section 2.

Remark. We note that the notation for the functions $f_{G,\mu,\alpha}$ and $\phi_{G,\mu,\alpha}$ is slightly abusive, as they also depend on the field E. Because confusion will not be possible, we have decided to drop the field E from the notations.

Proposition 4.2. *Let $P = MN$ be a standard parabolic subgroup of G. We have*

$$f^{(P)}_{G,\mu,\alpha} = q^{-\alpha\langle \rho_G - \rho_M, \mu\rangle} \sum_{w \in W_\alpha / \mathrm{stab}_{W_\alpha}(\mu) W_{M,\alpha}} f_{M,w(\mu),\alpha} \in \mathcal{H}_0(M),$$

where $\mathrm{stab}_{W_\alpha}(\mu) W_{M,\alpha} \subset W_\alpha$ *is the subgroup of W_α generated by the group $W_{M,\alpha}$ of the Weyl group of $T(E_\alpha)$ in $M(E_\alpha)$ and the stabilizer subgroup of μ in W_α.*

Proof. Compute the Satake transform of both sides to see that they are equal. $\square$

The integer α will later be the degree of the extension of the residue field of the reflex field at the prime of reduction over which we will count points in the Newton stratum. In this article we only want to show that the Newton-strata are non-empty. Therefore, we will take α large so that the combinatorial problems simplify (large in the divisible sense).

Let us make the function of Kottwitz more explicit in case G is either the restriction of scalars of a general linear group over F^+ or the restriction of scalars of an unramified unitary group over F^+. In fact, from this point onwards in this section we assume that we are in one of the following two cases:

$$G = \begin{cases} \mathrm{Res}_{F^+/\mathbb{Q}_p}(\mathrm{GL}_{n,F^+}) & \text{(linear type)} \\ \mathrm{Res}_{F^+/\mathbb{Q}_p}(U) & \text{(unitary type)} \end{cases} \tag{4.2}$$

where $F^+/\mathbb{Q}_p$ is a finite unramified extension, and where U/F^+ is an unramified unitary group, outer form of GL_{n,F^+}. Note that these groups G considered here occur as the components in the product decomposition in Equation (2.2). We assume that the cocharacter $\mu \in X_*(T)$ arises from a PEL-type datum, as we have explained in the discussion below Equation (2.1).

We begin with the linear case. We have a cocharacter $\mu \in X_*(T)$ (see below Equation (2.2)). Thus, for each $\mathbb{Q}_p$-embedding v of F^+ into $\overline{\mathbb{Q}}_p$ we get a cocharacter μ_v of the form

$$(\underbrace{1, 1, \ldots, 1}_{s_v}, \underbrace{0, 0, \ldots, 0}_{n-s_v}) \in \mathbb{Z}^n.$$

To each such integer s_v we associate the spherical function $f_{n\alpha s_v}$ on $\mathrm{GL}_n(F^+)$ whose Satake transform is defined by

$$\mathcal{S}(f_{n\alpha s_v}) = q^{\frac{s(n-s)}{2}\alpha} \sum_{i_1 < i_2 < \ldots < i_{s_v}} X^\alpha_{i_1} X^\alpha_{i_2} \cdots X^\alpha_{i_{s_v}} \in \mathbb{C}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]. \tag{4.3}$$

We write V_α for the set of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/E_\alpha)$-orbits in the set $\mathrm{Hom}(F^+, \overline{\mathbb{Q}}_p)$. If $v \in V_\alpha$ is such an orbit, then this orbit corresponds to a certain finite unramified extension $E_\alpha[v]$ of E_α. Let α_v be the degree over $\mathbb{Q}_p$ of the field $E_\alpha[v]$, we then have $E_\alpha[v] = E_{\alpha_v}$. The function f_α is given by

$$f_\alpha = \prod_{v \in V_\alpha} f_{n\alpha_v s_v}^{\mathrm{GL}_n(F^+)} \in \mathcal{H}_0(G(\mathbb{Q}_p)), \tag{4.4}$$

where the product is the convolution product (cf. [20, Prop. 3.3]).

Let us now assume that we are in the unitary case (cf. Equation (4.2)). We will make the function $f_{G,\mu,\alpha}$ explicit only in case α is even. To obtain the function of Kottwitz on G, we have to apply base change from $G(E_\alpha)$ to $G(\mathbb{Q}_p)$. Assume that α is even. Let $\mathbb{Q}_{p^2}$ be the quadratic unramified extension of $\mathbb{Q}_p$ contained in $\overline{\mathbb{Q}}_p$. The base change factors over the composition of base changes $G(E_\alpha) \rightsquigarrow G(\mathbb{Q}_{p^2}) \rightsquigarrow G(\mathbb{Q}_p)$. The base change of ϕ_α to $G(\mathbb{Q}_{p^2})$ is a function of the form $f_{G^+,\mu,(\alpha/2)}$ on the group $G^+ = \mathrm{Res}_{\mathbb{Q}_{p^2}/\mathbb{Q}_p}(G_{\mathbb{Q}_{p^2}})$. Explicitly, the last quadratic base change $G(\mathbb{Q}_{p^2}) \rightsquigarrow G(\mathbb{Q}_p)$ is given by:

$$\Psi \colon \mathbb{C}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]^{\mathfrak{S}_n} \longrightarrow \mathbb{C}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]^{\mathfrak{S}_m \rtimes (\mathbb{Z}/2\mathbb{Z})^m},$$

$$X_i \longmapsto \begin{cases} X_i & 1 \leq i \leq \lfloor n/2 \rfloor, \\ 1 & i = \lfloor \frac{n}{2} + 1 \rfloor, \text{ and } n \text{ is odd}, \\ X_{n+1-i}^{-1} & n+1-\lfloor n/2 \rfloor \leq i \leq n, \end{cases} \tag{4.5}$$

where $m := \lfloor \frac{n}{2} \rfloor$ (cf. [26]). Thus we get $f_{G,\mu,\alpha} = \Psi f_{G^+,\mu,(\alpha/2)}$.

Lemma 4.3. *Let G be an algebraic group over $\mathbb{Q}_p$ defined as in Equation (4.2). Let π be a generic unramified representation of G, and $f = f_{G,\mu,\alpha}$ a function of Kottwitz, and $b \in B(G)$ an isocrystal. Let $\alpha \in \mathbb{Z}_{>0}$ be an integer, sufficiently divisible such that W_α is the absolute the Weyl group of T in G. Then, the truncated trace $\mathrm{Tr}(\chi_b^G f_{G,\mu,\alpha}, \pi)$ is non-zero if and only if there exists some $w \in W_\alpha$ and some $\lambda \in \mathbb{R}_{>0}^\times$ such that $w(\mu) = \lambda \overline{\nu}_b \in \mathfrak{a}_0^G$.*

Remark. In case G is the general linear group, then there exists a pair $w \in W, \lambda \in \mathbb{R}_{>0}^\times$ such that $w(\mu) = \lambda \overline{\nu}_b$ if and only if the slopes λ_i of b all lie in the set $\{0, 1\}$.

Proof. We have $\pi = \mathrm{Ind}_T^G(\rho)$, where ρ is some smooth character of the torus T. By van Dijk's formula for truncated traces [20, Prop. 1.1] we have $\mathrm{Tr}(\chi_b^G f, \pi) = \mathrm{Tr}(\chi_b^G f^{(P_0)}, \rho)$. The truncation operation $h \mapsto \chi_b^G h$ on $\mathcal{H}_0(T)$, corresponds via the Satake transform to an operation on $\mathbb{C}[X_*(T)]$ which sends certain monomials $[M] \in \mathbb{C}[X_*(T)]$ associated to elements $M \in X_*(T)$, to zero and leaves certain other monomials invariant. Thus to compute the trace $\mathrm{Tr}(\chi_b^G f^{(P_0)}, \rho)$ one takes the set of monomials $[w(\mu)]$, $w \in W$ which occur in $f^{(P_0)}$, and removes some of them (maybe all), and then evaluate those that are left at the Hecke matrix of ρ. The first statement (i) of the Lemma now follows from the observation that $\chi_b^G \mathcal{S}_T^{-1}[w(\mu)] \neq 0$ if and only if $w(\mu) = \lambda \overline{\nu}_b$ for some positive scalar $\lambda \in \mathbb{R}_{>0}^\times$. This completes the proof. ∎

We have to distinguish further between (essentially) two cases at p. The case the group is the general linear group, and the case where the group is the unramified unitary group. We begin with the general linear group.

Proposition 4.4. *Let G be an algebraic group over $\mathbb{Q}_p$ defined as in Equation* (4.2)*, and assume it is of linear type, so $G(\mathbb{Q}_p) = \mathrm{GL}_n(F^+)$. Let $b \in B(G, \mu)$ be a μ-admissible isocrystal which has the property that the number of slopes equal to* 0 *is at most* 1*, and the number of slopes equal to* 1 *is also at most* 1*. Let χ be an unramified character of $\mathrm{GL}_n(F^+)$. Then, for α sufficiently divisible, we have $\mathrm{Tr}(\chi_b^G f_{G,\mu,\alpha}, \mathrm{St}_G(\chi)) \neq 0$.*

Remark. In the proof of the Proposition we use the divisibility of α at two places. First, it simplifies the function of Kottwitz (cf. Equation (4.4)). Second, we want α sufficiently divisible so that the Weyl group $W(T(E_\alpha), G(E_\alpha))$ relative to the field E_α is the full Weyl group.

Remark. In case the isocrystal b has two or more slopes with value 0 (or 1), then the truncated trace of the Kottwitz function on the Steinberg representation vanishes.

Proof. By Proposition 3.8 we have to show that the function $\xi_b^{\mathrm{St}} f_{G,\mu,\alpha}^{(P_0)}$ does not vanish. Recall that the function $f_{G,\mu,\alpha}$ is obtained from a function ϕ_α through base change from the group $\mathrm{GL}_n(F^+ \otimes E_\alpha)$. Let us first assume that the E_α-algebra $F^+ \otimes E_\alpha$ is a field. In that case we have that $f_{G,\mu,\alpha} = f_{n\alpha s}$ in the notations from [20, p. 10], i.e. $\mathcal{S}(f_{G,\mu,\alpha})$ is (up to scalar) an elementary symmetric function in the Satake algebra,

$$\mathcal{S}(f_{G,\mu,\alpha}) = q^{\frac{s_v(n-s_v)}{2}\alpha} \sum_{i_1 < i_2 < \ldots < i_s} X_{i_1}^{d\alpha} X_{i_2}^{d\alpha} \cdots X_{i_s}^{\alpha} \in \mathbb{C}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]. \tag{4.6}$$

We have to show that under the truncation operation $h \mapsto \xi_b^{\mathrm{St}} h$ on $\mathcal{H}(T)$ at least one of the monomials remains in Equation (4.6). Observe that the scalars in front of the monomials in Equation (4.6) all have the same sign, and that to get the truncated trace on the Steinberg representation we evaluate these monomials at a certain, nonzero point. Thus, the only problem is to see that there is at least one monomial X which occurs in $\mathcal{S}(f_{G,\mu,\alpha})$ and which survives the truncation $X \mapsto \xi_b^{\mathrm{St}} X$. At this point it will be useful to give a graphical interpretation of this truncation process.

A remark on the notation: With $\xi_b^{\mathrm{St}} X$ for X a monomial in the Satake algebra of T, we mean the element $\mathcal{S}_T(\xi_b^{\mathrm{St}} \mathcal{S}_T^{-1}(X))$ of the Satake algebra of T. Below we will use similar conventions for the truncations $\chi_N X$, $\widehat{\chi}_{N_0 \cap M_b} X$ and $\eta_b X$.

A *graph* in $\mathbb{Z}^2$ is a sequence of points $\vec{v}_0, \vec{v}_1, \ldots, \vec{v}_r$ with $\vec{v}_{i+1} - \vec{v}_i = (1, e)$, where e is an integer. To a monomial $X = X_1^{e_1} X_2^{e_2} \cdots X_n^{e_n} \in \mathbb{C}[X_1^{\pm 1}, X_2^{\pm 1}, \ldots, X_n^{\pm 1}]$, with $e_i \in \mathbb{Z}$ and $\sum_{i=1}^n e_i = s$ we associate the graph $\mathcal{G}_X$ with points

$$\vec{v}_0 := (0,0), \quad \vec{v}_i := \vec{v}_0 + (i, e_n + e_{n-1} + \ldots + e_{n+1-i}) \in \mathbb{Z}^2, \tag{4.7}$$

for $i = 1, \dots, n$. Because the sum $\sum_{i=1}^n e_i$ is equal to s, we see that the end point of the graph is (n, s). The function $f_{G\alpha\mu}$ is (up to scalar) the elementary symmetric function of degree s in n variables, thus its monomials correspond precisely to the set of graphs that start at the point $(0, 0)$, have end point (n, s) and satisfy $\vec{v}_{i+1} - \vec{v}_i \in \{(1, 0), (1, 1)\}$ for all i.

To the slopes $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ of the isocrystal b we associate the graph $\mathcal{G}_b$ with points

$$\vec{v}_0 := (0, 0), \quad \vec{v}_i := \vec{v}_0 + (i, \lambda_1 + \lambda_2 + \ldots + \lambda_i) \in \mathbb{Z}^2, \tag{4.8}$$

for $i = 1, \dots, n$. (Remark: To obtain the usual convex picture of the Newton polygon we had to invert the order of the vector $e_1, \dots, e_n$ in Equation (4.7). Without the inversion we would be considering concave polygons.)

We may now explain the truncation $X \mapsto \xi_b^{St} X$ in terms of graphs. We have $\xi_b^{\mathrm{St}} X = X$ or $\xi_b^{\mathrm{St}} X = 0$. We claim that we have $\xi_b^{\mathrm{St}} X = X$ if the following conditions hold:

(**C1**) We have $\mathcal{G}_b(n) = \lambda \mathcal{G}_X(n)$ for some positive scalar $\lambda \in \mathbb{R}_{>0}$;
(**C2**) For every break point $x \in \mathbb{Z}^2$ of $\mathcal{G}_b$ the point x also lies on the graph $\lambda \mathcal{G}_X$;
(**C3**) Outside the set of breakpoints of $\mathcal{G}_b$, the graph $\lambda \mathcal{G}_X$ lies strictly below the graph $\mathcal{G}_b$.

Thus, in short: $\mathcal{G}_X$ lies below $\mathcal{G}_b$ and the set of contact points between the two graphs is precisely the begin point, end point and the set of break points of $\mathcal{G}_b$. See also our preprint [21], where we use this construction in an analogous context.

Remark. In the claim above we say "if" and not "if and only if". The conditions (**C1**), (**C2**) and (**C3**) are stronger than the condition $\xi_b^{\mathrm{St}} X = X$. In Lemmas 4.5, 4.6 and 4.7 below we give conditions (**C1'**), (**C2'**) and (**C3'**) which, when taken together, are equivalent to "$\xi_b^{\mathrm{St}} X = X$". However (**C1**, **C2**, **C3**) is not equivalent to (**C1'**, **C2'**, **C3'**). If one would one replace condition (**C3'**) with the stronger condition

$$(\textbf{C3"}) \quad \text{We have } \overline{\mathcal{G}}_x = \lambda \mathcal{G}_b \text{ for some } \lambda \in \mathbb{R}_{>0},$$

then we have $(\textbf{C1}, \textbf{C2}, \textbf{C3}) \iff (\textbf{C1'}, \textbf{C2'}, \textbf{C3"})$.

Because the above fact is crucial for the argument, let us prove the claim with complete details. Let $X = (e_1, e_2, \dots, e_n) \in \mathbb{Z}^n = X_*(A_0)$. We want to express the condition $\xi_b^{\mathrm{St}} X = X$ in terms of $\mathcal{G}_X$. Recall that the Satake transform for the maximal torus $T = (\mathrm{Res}_{F^+/\mathbb{Q}_p} \mathbb{G}_{\mathrm{m}})^n$ is simply

$$\begin{aligned} \mathcal{H}_0(T) &\xrightarrow{\sim} \mathbb{C}[X_1^{\pm 1}, X_2^{\pm 1}, \dots, X_n^{\pm 1}], \\ \mathbf{1}_{(p^{-e_1}\mathcal{O}_{F^+}^\times) \times (p^{-e_2}\mathcal{O}_{F^+}^\times) \times \cdots \times (p^{-e_n}\mathcal{O}_{F^+}^\times)} &\longmapsto X^{e_1} X^{e_2} \cdots X^{e_n}. \end{aligned} \tag{4.9}$$

We have $\xi_b^{\mathrm{St}} = \widehat{\chi}_{N_0 \cap M_b} \chi_{N_b} \eta_b$. Let (n_a) be the composition of n corresponding to the standard parabolic subgroup P_b of G. Let $g = (g_1, \dots, g_n) \in T$ such that $\chi_{N_b}(g) = 1$. Explicitly, this means that

$$|g_1 g_2 \cdots g_{n_1}|^{1/n_1} < |g_{n_1+1} g_{n_1+2} \cdots g_{n_1+n_2}|^{1/n_2} < \ldots < |g_{n_{k-1}} g_{n_{k-1}+1} \cdots g_n|^{1/n_k}, \tag{4.10}$$

(cf. [20, Eq. (1.11)]). In terms of the graph $\mathcal{G}_X$ of X this means the following. We have $X \in \mathfrak{a}_0$ and we have the projection $\pi_{0,P_b}(X)$ of X in $\mathfrak{a}_{P_b}$ (obtained by taking the average under the action of the Weyl group of M_b). We write $\overline{\mathcal{G}}_X$ for the graph of $\pi_{0,P_b}(X) \in \mathfrak{a}_{P_b} \subset \mathfrak{a}_0$. This graph $\overline{\mathcal{G}}_X$ is obtained from the graph $\mathcal{G}_X$ as follows. Consider the list of points

(4.11)
$$p_0 := (0,0), \quad p_1 := (n_1, \mathcal{G}_X(n_1)), \quad p_2 := (n_1+n_2, \mathcal{G}_X(n_1+n_2)), \quad \ldots \quad p_k := (n, \mathcal{G}_X(n)).$$

Connect, using a straight line, the point p_0 with p_1, and with another straight line, the point p_1 with p_2, etc, to obtain the graph $\overline{\mathcal{G}}_X$ from $\mathcal{G}_X$. From Equations (4.9) and (4.10) we get:

Lemma 4.5. *For a monomial X we have $\chi_{N_b} X = X$ if and only if the following condition is true:*

(C1') The graph $\overline{\mathcal{G}}_X$ is convex.

(Remark: We have $\chi_{N_b} X = 0$ if condition (**C1'**) is not satisfied. This remark also applies to Lemmas 4.6 and 4.7.)

Before discussing the function $\widehat{\chi}_{N_0 \cap M_b}$, let us first discuss in detail the maximal case, i.e. the function $\widehat{\chi}_{N_0}$ for the group G (cf. [20, Prop. 1.11]). We have $\mathfrak{a}_0 = \mathbb{R}^n$, write $H_1, \ldots, H_n$ for the basis of $\mathfrak{a}_0^*$ dual to the standard basis of $\mathbb{R}^n$. Write α_i for root $H_i - H_{i+1}$ in $\mathfrak{a}_0^*$. We have

$$\varpi_{\alpha_i}^G = \left(H_1 + H_2 + \ldots + H_i - \frac{i}{n} (H_1 + H_2 + \ldots + H_n) \right) \in \mathfrak{a}_0^{G*}. \tag{4.12}$$

Thus, for a monomial $X = X_1^{e_1} X_2^{e_2} \cdots X_n^{e_n}$ the condition $\langle \varpi_{\alpha_i}^G, X \rangle > 0$ corresponds to

$$e_1 + e_2 + \ldots + e_i > \frac{i}{n} (e_1 + e_2 + \ldots + e_n) \tag{4.13}$$

Thus we obtain

$$\mathcal{G}_X(n+1-i) > \frac{i}{n} s, \tag{4.14}$$

where s is the degree of X, i.e. $s = \sum_{i=1}^n e_i$. If we ask that $\langle \varpi_\alpha^G, X \rangle > 0$ holds for all roots α of G, then this means precisely that the graph $\mathcal{G}_X$ lies strictly below[2] the straight line which connects the point $(0,0)$ with the point $(n, \mathcal{G}_X(n))$.

We now turn to the function $\widehat{\chi}_{N_0 \cap M_b}$. The group M_b decomposes into a product of general linear groups, say it corresponds to the composition (n_a) of the integer n. Thus, the condition

$$\forall \alpha \in \Delta^{M_b} : \langle \varpi_\alpha^{M_b}, X \rangle > 0,$$

is the condition in Equation (4.13) but, then for each of the blocks of M_b individually. The conclusion is that:

Lemma 4.6. *For any monomial X we have $\widehat{\chi}_{N_0 \cap X_b} \cdot X = X$ if and only if the following condition is true:*

(C2') The graph $\mathcal{G}_X$ lies below $\overline{\mathcal{G}}_X$ and the two graphs touch precisely at the points p_i.

[2]We get 'below' and not 'above' due to the inversion "$e_i \mapsto e_{n+1-i}$" in Equation (4.7).

The condition that $\eta_b X = X$ means that $\pi_{0,P_b}(\Phi(g)) = \lambda\overline{\nu}_b$ whenever g lies in the support of the function $\mathcal{S}_T^{-1}(X)$ on the group T. By the explicit formula for the Satake transform (Equation (4.9)), the condition is equivalent to the condition that there exists a $w \in \mathfrak{S}_n$ such that the vector

$$\left(\underbrace{e_{w(1)} + e_{w(2)} + \cdots}_{n_1}, \underbrace{e_{w(n_1+1)} + e_{w(n_1+2)} + \cdots}_{n_2}, \ldots, \underbrace{e_{w(n_1+n_2+\ldots+n_{k-1}+1)} + \cdots}_{n_k}\right) \in \mathfrak{a}_{P_b},$$

is a positive scalar multiple of the vector $\overline{\nu}_b$. Using earlier notations, this can be expressed as:

Lemma 4.7. *For any monomial X we have $\eta_b X = X$ if and only if the following condition is true:*

(**C3'**) There exists an element $w \in \mathfrak{S}_n$ such that $\overline{\mathcal{G}}_{w(X)} = \lambda\mathcal{G}_b$ for some $\lambda \in \mathbb{R}_{>0}$.

To prove the claim we have to show that the group of conditions (**C1**), (**C2**) and (**C3**) implies the group of conditions (**C1'**), (**C2'**) and (**C3'**).

Thus, assume the conditions (**C1**), (**C2**) and (**C3**) are true for the monomial X. Recall that P_b is the parabolic subgroup contracted by the isocrystal b. This means that the set of breakpoints of the polygon $\mathcal{G}_b$ is equal to the set

$$q_0 = (0,0), \quad q_1 = (n_1, \mathcal{G}_b(n_1)), \quad q_2 = (n_1+n_2, \mathcal{G}_b(n_1+n_2)) \quad \ldots \quad q_k = (n, \mathcal{G}_b(n)).$$

By condition (**C1**) there is a $\lambda \in \mathbb{R}_{>0}$ such that $\mathcal{G}_b(n) = \lambda\mathcal{G}_X(n)$. By conditions (**C2**) and (**C3**) the set $\{q_0, \ldots, q_n\}$ is then precisely the set of points where the graph $\lambda\mathcal{G}_X$ touches the graph $\mathcal{G}_b$. Taking averages, we get the relation $\overline{\mathcal{G}}_b = \lambda\overline{\mathcal{G}}_X$. We have $\mathcal{G}_b = \overline{\mathcal{G}}_b$ (because P_b is associated to b), and therefore $\mathcal{G}_b = \lambda\overline{\mathcal{G}}_X$. Thus condition (**C3'**) is true for $w = \mathrm{Id} \in \mathfrak{S}_n$. The condition (**C2'**) is now implied by (**C2**) and (**C3**). Finally let us prove condition (**C1'**). We have $\lambda\overline{\mathcal{G}}_X = \mathcal{G}_b$, and the graph $\mathcal{G}_b$ is convex. Thus $\overline{\mathcal{G}}_X$ is convex. The three conditions (**C1'**), (**C2'**) and (**C3'**) are now verified, and therefore the claim is true.

Recall that the monomials M which occur in $\mathcal{S}(f_{G,\mu,\alpha})$ corresponds to the set of graphs from $(0,0)$ to (n,s) whose steps consist of diagonal, north-eastward steps, or horizontal, eastward steps. Thus, it suffices to show that there exists a graph which satisfies the conditions (**C1**), (**C2**) and (**C3**) above. This is indeed possible under the condition on the slopes of λ_i of b (see Figure 2 for the explanation). This completes the proof in case $F^+ \otimes E_\alpha$ is a field.

We now drop the assumption that $F^+ \otimes E_\alpha$ is a field. By [20, Prop. 3.3] there exists a sufficiently large integer $M \geq 1$ such that for all degrees α which are divisible by M, the function $f_{G\mu\alpha}$ is (up to a scalar) a convolution product of the form $\prod_{i=1}^r f_{n\alpha s_i}$, where $r = [F^+ : \mathbb{Q}_p]$ and (s_i) is a certain given composition of an integer s of length r. Observe that any monomial which occurs in $\mathcal{S}(f_{n\alpha s})$ also occurs in the product $\prod_{i=1}^r \mathcal{S}(f_{n\alpha s_i})$ with a positive coefficient. Thus we may write $\prod_{i=1}^r f_{n\alpha s_i} = f_{n\alpha s} + R \in \mathcal{H}(G)$ for some function $R \in \mathcal{H}(G)$, whose Satake transform is a linear combination of monomials, with all coefficients

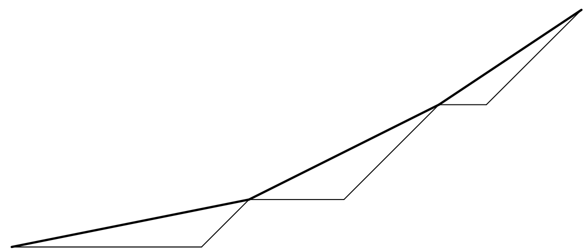

FIGURE 2. The dark line is an example of the Newton polygon of an isocrystal b with additional $\mathrm{GL}_{12}(F^+)$-structure whose slope morphism is $(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{2}{3}, \frac{2}{3}, \frac{2}{3})$. The thin line is a ξ_b^{St}-admissible path. One can see that for this Newton polygon there exist precisely two admissible paths. In general one can always take the 'ordinary' path which, within the blocks where the Newton polygon is of constant slope, starts with horizontal steps, and then ends with diagonal steps. Observe that however in case there are slopes equal to 0 in $\overline{\nu}_b$, then in the slope 0 block there exists a unique path $\mathcal{G}$ which connects the start point with the end point. Furthermore, when the slope 0 block is of size strictly larger than 1, then on the path $\mathcal{G}$ there is a point which does not lie strictly below $\mathcal{G}_b$, and this violates the ξ_b^{St}-condition. The same problem occurs if there are > 1 slopes equal to 1.

positive. Consequently, to check that the truncated trace of $\prod_{i=1}^r f_{n\alpha s_i}$ on the Steinberg representation is non-zero, it suffices to check that the truncated trace of $f_{n\alpha s}$ on Steinberg is non-zero. This completes the proof. $\square$

Proposition 4.8. *Let G be an algebraic group over $\mathbb{Q}_p$ defined as in Equation (4.2), and assume it is of linear type. Let $b \in B(G, \mu)$ be a μ-admissible isocrystal. Let m_0 be the number of indices i such that $\lambda_i = 0$, and let m_1 be the number of indices i such that $\lambda_i = 1$. Write $m := n - m_0 - m_1$. Let π_{m_0} (resp. π_{m_1}) be any generic unramified representation of $\mathrm{GL}_{m_0}(F^+)$ (resp. $\mathrm{GL}_{m_1}(F^+)$), and χ an unramified character of $\mathrm{GL}_m(F^+)$. Let P be the standard parabolic subgroup of G with 3 blocks, the first of size m_1, the second of size m and the last one of size m_3. Then for α sufficiently divisible we have*

$$\mathrm{Tr}\left(\chi_b^G f_{G,\mu,\alpha}, \mathrm{Ind}_P^G\left(\pi_{m_1} \otimes \mathrm{St}_{\mathrm{GL}_m(F^+)}(\chi) \otimes \pi_{m_0}\right)\right) \neq 0.$$

Remark. Note that we have abused language slightly when we said that P has 3 parabolic subgroups. We could have m, m_0 or m_1 equal to 0, in which case P has less than 3 blocks. If one of the numbers m, m_0 or m_1 is 0, then one simply removes the corresponding factor from tensor product $\pi_{m_1} \otimes \mathrm{St}_{\mathrm{GL}_m(F^+)}(\chi) \otimes \pi_{m_0}$, and one induces from a parabolic subgroup with two blocks (or one block).

Proof of Proposition 4.8. By van Dijk's formula for truncated traces [20, Prop. 1.5], we get a trace on M:

$$\mathrm{Tr}\left(\chi_b^G f_{G,\mu,\alpha}^{(P)}, \pi_{m_1} \otimes \mathrm{St}_{\mathrm{GL}_m(F^+)} \otimes \pi_{m_0}\right). \tag{4.15}$$

By Proposition 4.2 we have

$$f_{G,\mu,\alpha}^{(P)} = q^{-\alpha\langle \rho_G - \rho_M, \mu\rangle} \sum_{w \in W_{\alpha})/\mathrm{stab}_{W_{\alpha})}(\mu) W_{M,\alpha}} f_{M,w(\mu),\alpha} \in \mathcal{H}_0(M). \tag{4.16}$$

The intersection $\Omega_{\overline{\nu}_b}^G \cap M$ is equal to a union $\bigcup \Omega_{w(\overline{\nu}_b)}^M$, where w ranges over the subset of elements $w \in W$ such that $w(\overline{\nu}_b) \in \mathfrak{a}_0$ is M-positive. Consequently, if we plug Equation (4.16) into Equation (4.15), then we get a large sum, call it $(\star)$, of traces of functions $f_{M,w(\mu),\alpha}$ against a representation of the form $\pi_{m_1} \otimes \mathrm{St}_{\mathrm{GL}_m(F^+)} \otimes \pi_{m_0}$. All the signs are the same in this large sum $(\star)$, therefore it suffices to see that there is at least one non-zero term. Take $b_M \in B(M)$ the isocrystal whose slope morphism is $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ in the M-positive chamber of $\mathfrak{a}_0$. Then b_M has only slopes 0 on the first block of M and only slopes 1 on the third block, and all its slopes $\neq 0, 1$ are in the second block. The trace $\mathrm{Tr}(\chi_{b_M}^M f_{M,\mu,\alpha}, \pi_{m_1} \otimes \mathrm{St}_{\mathrm{GL}_m(F^+)} \otimes \pi_{m_0})$ occurs as a term in the expression $(\star)$. By Lemma 4.3 and Proposition 4.4 this term is non-zero. This completes the proof. □

We now establish the cases where the group is an unramified unitary group over F^+ (unitary type, cf. Equation (4.2)).

Lemma 4.9. *Let G be an algebraic group over $\mathbb{Q}_p$ defined as in Equation* (4.2)*, and assume it is of unitary type. Assume that $b \in B(G,\mu)$ is an isocrystal whose slope morphism $\overline{\nu}_b \in \mathfrak{a}_0$ has no coordinate equal to* 0 *and no coordinate equal to* 1*. Then, for α sufficiently divisible, the trace $\mathrm{Tr}(\chi_b^G f_{G,\mu,\alpha}, \mathrm{St}_{G(\mathbb{Q}_p)})$ is non-zero.*

Proof. We use the explicit description $f_{G,\mu,\alpha} = \Psi f_{G^+,\mu,(\alpha/2)}$ of the Kottwitz function that we gave in Equation (4.5). Let us assume that the algebra $F^+ \otimes E_\alpha$ is a field; then the base change mapping from $G(F_\alpha^+) \to G(F_2)$ is given by $X_i \mapsto X_i^{\alpha/2}$ on the Satake algebras. Over F_α^+, the Weyl group W_α is equal to $\mathfrak{S}_n$ with its natural action on $\mathbb{R}^n$. Note that the formula for the base change mapping Ψ from Equation (4.5) also makes sense over the Satake algebras of the maximal split tori, i.e. we have a map Ψ from the algebra $\mathbb{C}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]$ to $\mathbb{C}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]$. The monomials which occur in $f_{G,\mu,\alpha}$ are those monomials which are of the form $\Psi[w(\mu)]$ where w is some element of $\mathfrak{S}_n$. The Weyl group translates $[w(\mu)]$ of $[\mu]$ correspond to all paths from $(0,0)$ to (n,s), and the monomials of the form $\Psi[w(\mu)] = [w(\mu)] + [\theta(w\mu)]$ correspond to all paths from $(0,0)$ to $(n,0)$ which stay below the horizontal line with equation $y = s$, and above the horizontal line with equation $y = -s$. We have that $\chi_b^{G(\mathbb{Q}_p)} \Psi[w(\mu)] \neq 0$ if the path $\mathcal{G}$ of $\Psi[w(\mu)]$ lies below $\mathcal{G}_b$ and the set of contact points between the two graphs is precisely the begin point, end point and the set of break points of $\mathcal{G}_b$. This is the same condition as had for the general linear group (see above Equation (4.9))

because the root systems are the same. Again, by drawing a picture, it is easy to see that such graphs exist in case b has no slopes equal to $-1, 0$ or 1. Consequently $\chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha} \neq 0$, and then also $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}, \mathrm{St}_G) \neq 0$ by Proposition 3.8.

If we forget the assumption that $F^+ \otimes E_\alpha$ is a field, then, we proceed just as we did for the general linear group (cf. Lemma 4.4), we write $f_{G,\mu,\alpha} = A + R$, where R is a function whose Satake transform is a linear combination of monomials in the Satake algebra with all coefficients positive, and A is a function for which we already know that its truncated trace on the Steinberg representation does not vanish. This completes the proof. $\square$

Proposition 4.10. *Let G be an algebraic group over $\mathbb{Q}_p$ defined as in Equation* (4.2)*, and assume it is of unitary type. Let $b \in B(G, \mu)$ be an isocrystal with slopes $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ (cf. the discussion below Proposition 1.1). Let $n = m_1 + m_2 + m_3$ be the composition of n with the property that the first block of m_1 slopes λ_i satisfy $\lambda_i = -1$, the second block of slopes satisfy $-1 < \lambda_i < 1$ and is of size m_2, the third block of slopes satisfy $\lambda_i = 1$ and is of size m_3. We have $m_1 = m_3$. Let $P = MN$ be the standard parabolic subgroup of G corresponding to this composition of n, thus M is a product of two groups, $M = M_1 \times M_2$, where $M_1 = \mathrm{GL}_{m_1}(F^+)$ is a general linear group and M_2 is an unramified unitary group. For α sufficiently divisible the trace $\mathrm{Tr}(\chi_b^{G(F^+)} f_{G,\mu,\alpha}, \bullet)$ against the representation $\mathrm{Ind}_{P(F^+)}^{G(F^+)}(\pi_{m_1} \otimes \mathrm{St}_{m_2}(\chi))$ is non-zero if π_{m_1} is an unramified generic representation and χ an unramified character of $\mathrm{GL}_{m_2}(F^+)$.*

Remark. We could have that M_1 is the trivial group, this happens in case $-1 < \lambda < 1$ for all indices i. In that case the considered representation is simply an unramified twist of the Steinberg representation.

Proof. The proof is the same as the proof in case of the general linear group (cf. Proposition 4.8): one easily reduces the statement to that of Lemma 4.9. $\square$

Let now $G/\mathbb{Q}$ be an unitary group of similitudes which arises from a Shimura datum of PEL-type (cf. Equation (2.1)), and let $G_1 \subset G$ be the kernel of the factor of similitudes. The group G_1 is defined over a totally real field F^+, and defined with respect to a quadratic extension F of F^+, which is a CM field. Let $A_0 \subset G$ be a maximally split torus, then we may write $A_0 = \mathbb{G}_{\mathrm{m}} \times A_0'$ (*not* a direct product), where $A_0' \subset G_1$ be the maximally split torus of G_1 defined by $G_1 \cap A_0$. At p we have a decomposition of $F^+ \otimes \mathbb{Q}_p$ into a product of fields $F_\wp^+$, where $\wp$ ranges over the primes above p. Let p be a prime number where G is unramified. The group $G_{1,\mathbb{Q}_p}$ is of the form $G_{1,\mathbb{Q}_p} \cong \prod_\wp \mathrm{Res}_{F_\wp^+/\mathbb{Q}_p} G_{1,\wp}$, where the group $G_{1,\wp}$ is either an unramified unitary group over $F_\wp^+$, or the general linear group. In the first case we call the F^+-prime $\wp$ *unitary* and in the second case we call the prime *linear*.

Consider an isocrystal $b \in B(G)$. To b we may associate its slope morphism $\overline{\nu}_b \in \mathfrak{a}_0$. Let $A_{0,\wp}' \subset G_{1,\wp}$ be the $\wp$-th component of A_0'; it is a split maximal torus in $G_{1,\wp}$, and write $\mathfrak{a}_0(\wp) := X_*(A_{0,\wp}')$. The space $\mathfrak{a}_0$ decomposes along the split center and the F^+-primes $\wp$ above p: $\mathfrak{a}_0 = \mathbb{R} \times \prod_\wp \mathfrak{a}_0(\wp)$. Thus we can speak for each $\wp$ of the $\wp$-component $\overline{\nu}_{b,\wp}$ of $\overline{\nu}_b$.

In case $\wp$ is linear, the Proposition 4.8 gives us a class of representations $\pi'_\wp$ of $G_{1,\wp}(\mathbb{Q}_p)$ such that the $b_\wp$-truncated trace on $\pi'_\wp$ does not vanish. In case $\wp$ is unitary, we get such a class $\pi'_\wp$ from Proposition 4.10. Let π' be the representation of $G_1(\mathbb{Q}_p)$ obtained from the factors $\pi_\wp$ by taking the tensor product.

Definition 4.11. We write $\mathfrak{R}_1(b)$ for the just constructed class of $G_1(\mathbb{Q}_p)$-representations π'.

Observe that the set of representations $\mathfrak{R}_1(b)$ has positive Plancherel measure in the set of $G_1(\mathbb{Q}_p)$ representations, and that the b-truncated trace of the Kottwitz function on these representations does not vanish by construction.

We now extend the class $\mathfrak{R}_1(b)$ to a class of $G(\mathbb{Q}_p)$-representations, as follows:

Definition 4.12. Let $\pi \in \mathfrak{R}_1(b)$. Then π is an $H(\mathbb{Q}_p)$-representation; let ω_π be its central character, thus ω_π is a character of $Z_1(\mathbb{Q}_p)$. Assume χ is a character of $Z(\mathbb{Q}_p)$ which extends ω_π. Then we may extend the representation π to a representation $\pi\chi$ of the group $H(\mathbb{Q}_p)Z(\mathbb{Q}_p)$. We define $\mathfrak{R}_1(b)'$ to be the set of $H(\mathbb{Q}_p)Z(\mathbb{Q}_p)$-representations of the form $\pi\chi$. Notice that not all the inductions $\mathrm{Ind}_{H(\mathbb{Q}_p)}^{G(\mathbb{Q}_p)}(\pi\chi)$ have to be irreducible, we will ignore those which are reducible. We define $\mathfrak{R}(b)$ to be the set of representations Π which are isomorphic to an irreducible induction $\mathrm{Ind}_{H(\mathbb{Q}_p)Z(\mathbb{Q}_p)}^{G(\mathbb{Q}_p)}(\pi\chi)$ with $\pi\chi \in \mathfrak{R}_1(b)'$.

The required non-vanishing property of the representations in $\mathfrak{R}_b$ will be shown in the next section.

5. Local extension

We need to extend from $G_1(\mathbb{Q}_p)$ to the group $G(\mathbb{Q}_p)$. Let Z be the center of the group G. Consider the morphism of algebraic groups $\psi\colon G_{1,\mathbb{Q}_p} \times Z_{\mathbb{Q}_p} \twoheadrightarrow G_{\mathbb{Q}_p}$; the group $\mathrm{Ker}(\psi)$ is the center Z_1 of the group G_1, so

$$\mathrm{Ker}(\psi) = \prod_\wp \begin{cases} \mathbb{G}_{\mathrm{m}} & \wp \text{ is linear} \\ U_1^* & \wp \text{ is unitary,} \end{cases} \tag{5.1}$$

where U_1^* is the unramified non-split form of $\mathbb{G}_{\mathrm{m}}$ over $F_\wp^+$. Over $\mathbb{Q}$, Z is defined by $Z(\mathbb{Q}) = \{x \in F^\times | N_{F/F^+}(x) \in \mathbb{Q}^\times\}$. Using Equation (5.1), the long exact sequence for Galois cohomology and Shapiro's lemma, the group $G(\mathbb{Q}_p)/G_1(\mathbb{Q}_p)Z(\mathbb{Q}_p)$ maps injectively into the group $(\mathbb{Z}/2\mathbb{Z})^t$, where t is the number of unitary places of F^+ above p.

Write $\mu' \in X_*(T)$ for the cocharacter of the maximal torus $(T \cap G_1) \cap Z$ of $G_1 \times Z$ obtained from μ via restriction. Let $f_{G_1 \times Z, \mu', \alpha}$ be the corresponding function of Kottwitz on the group $G_1(\mathbb{Q}_p) \times Z(\mathbb{Q}_p)$. Furthermore we write $\chi_b^{G_1 \times Z}$ for the characteristic function on $G_1(\mathbb{Q}_p) \times Z(\mathbb{Q}_p)$ of elements (g, z) such that we have $\chi_b^{G(\mathbb{Q}_p)}(gz) = 1$. We prove the following statement:

Proposition 5.1. *Fix a representation π_0 of $G_1(\mathbb{Q}_p)$. Let Π be a smooth irreducible representation of $G(\mathbb{Q}_p)$, which, upon restriction to $G_1(\mathbb{Q}_p) \times Z(\mathbb{Q}_p)$, contains the representation*

π_0 of $G_1(\mathbb{Q}_p)$. Assume the central character of Π is of finite order. Then, for all sufficiently divisible α, we have

$$\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}, \Pi) = t(\Pi) \, \mathrm{Tr}(\chi_b^{G_1 \times Z} f_{G_1 \times Z,\mu,\alpha}, \pi_0)$$

where $t(\Pi)$ is a positive real number.

Before proving Proposition 5.1 we first establish some technical results. We fix smooth models of G, G_1, Z, etc. over $\mathbb{Z}_p$ (and use the same letter for them). We have the exact sequence $Z_1 \rightarrowtail Z \times G_1 \twoheadrightarrow G$, so the cokernel of $Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p)$ in $G(\mathbb{Q}_p)$ is a subgroup of $H^1(\mathbb{Q}_p, Z_1) \cong (\mathbb{Z}/2\mathbb{Z})^t$, where t is the number of unitary places.

Lemma 5.2. *The mapping $G_1(\mathbb{Z}_p) \times Z(\mathbb{Z}_p) \to G(\mathbb{Z}_p)$ is surjective.*

Proof. We have $Z_1 \rightarrowtail G_1 \times Z \twoheadrightarrow G$, which is an exact sequence of algebraic groups over $\mathrm{Spec}(\mathbb{Z}_p)$. Thus we get $Z_1(\mathbb{F}_p) \rightarrowtail G_1(\mathbb{F}_p) \times Z(\mathbb{F}_p) \to G(\mathbb{F}_p) \to H^1(\mathbb{F}_p, Z_1)$. The group Z_1 is a torus and therefore connected. By Lang's theorem we obtain $H^1(\mathbb{F}_p, Z_1) = 1$. Thus the mapping $G_1(\mathbb{F}_p) \times Z(\mathbb{F}_p) \to G(\mathbb{F}_p)$ is surjective. By Hensel's lemma the mapping $G_1(\mathbb{Z}_p) \times Z(\mathbb{Z}_p) \to G(\mathbb{Z}_p)$ is then also surjective. □

Lemma 5.3. *The function of Kottwitz $f_{G,\mu,\alpha}$ has support on the subset $Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p) \subset G(\mathbb{Q}_p)$.*

Proof. Define χ on $G(\mathbb{Q}_p)$ to be the characteristic function of the subset $Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p) \subset G(\mathbb{Q}_p)$. The mapping $Z \times G_1 \to G$ is surjective on $\mathbb{Z}_p$-points, and therefore χ is spherical. The functions $\chi f_{G,\mu,\alpha}^G$ and $f_{G,\mu,\alpha}$ are then both spherical functions and to show that they are equal it suffice to show that their Satake transforms agree (the Satake transform is injective). We have $\mathcal{S}(\chi f_{G,\mu,\alpha}) = \chi|_{A(\mathbb{Q}_p)} \mathcal{S}(f_{G,\mu,\alpha})$, where A is a maximal split torus of G, $\chi|_{A(\mathbb{Q}_p)}$ is the characteristic function of the subset $Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p) \cap A(\mathbb{Q}_p) \subset A(\mathbb{Q}_p)$. Observe that, in fact, $Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p) \cap A(\mathbb{Q}_p) = A(\mathbb{Q}_p)$. This implies $\chi|_{A(\mathbb{Q}_p)} \mathcal{S}(f_{G,\mu,\alpha}) = \mathcal{S}(f_{G,\mu,\alpha})$, and shows that $\chi f_{G,\mu,\alpha}$ and $f_{G,\mu,\alpha}$ have the same Satake transform. This completes the proof of the lemma. □

We now turn to the proof of Proposition 5.1.

Proof of Proposition 5.1. By Clifford theory [35, thm 2.40] the representation Π restricted to $G_1(\mathbb{Q}_p)Z(\mathbb{Q}_p)$ is a finite direct sum of irreducible representations π_i, where π_i satisfies $\pi_i(g) = \pi_0(x_i g x_i^{-1})$ for some x_i not depending on g. We clarify that in this finite direct sum multiplicities may occur. As characters on $G_1(\mathbb{Q}_p)Z(\mathbb{Q}_p)$ we may write $\theta_\Pi = \sum_{i=1}^t \theta_{\pi_i} \omega_i$, where θ_{π_i} is the Harish-Chandra character of π_i, viewed as a $G_1(\mathbb{Q}_p)$-representation, and ω_i

is the central character of π_i. Using Lemma 5.3 we may now compute:

$$\begin{aligned}
\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}, \Pi) &= \int_{Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p)} \chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha} \theta_\Pi \mathrm{d}g \\
&= \sum_{i=1}^{t} \int_{Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p)} \chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha} \theta_{\pi_i} \omega_i \mathrm{d}g \\
&= \sum_{i=1}^{t} \int_{Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p)} \chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}^{x_i^{-1}} \theta_{\pi_0} \omega_0 \mathrm{d}g, \qquad (5.2)
\end{aligned}$$

where $f_{G,\mu,\alpha}^{x_i^{-1}}$ is the conjugate of $f_{G,\mu,\alpha}$ by x_i^{-1}. Note, however, that the function of Kottwitz is stable under the action of the Weyl group of G. Therefore $f_{G,\mu,\alpha}^{x_i^{-1}} = f_{G,\mu,\alpha}$. We get the expression:

$$t \int_{Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p)} \chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha} \theta_{\pi_0} \omega_0 \mathrm{d}g.$$

On the other hand we have

$$0 \neq \mathrm{Tr}(\chi_b^{Z \times G_1} f_{Z \times G_1, \mu', \alpha}, \pi_0) = \int_{Z(\mathbb{Q}_p) \times G_1(\mathbb{Q}_p)} \chi_b^{Z \times G_1} f_{Z \times G_1, \mu', \alpha} [\theta_{\pi_0} \times \omega_0] \mathrm{d}g.$$

We compute the right hand side:

$$\begin{aligned}
&\int_{\frac{Z(\mathbb{Q}_p) \times G_1(\mathbb{Q}_p)}{Z_1(\mathbb{Q}_p)}} \int_{Z_1(\mathbb{Q}_p)} (\chi_b^{Z \times G_1} f_{Z \times G_1, \mu', \alpha} [\theta_{\pi_0} \times \omega_0])(zz_1, hz_1) \mathrm{d}z_1 \frac{\mathrm{d}(z,h)}{\mathrm{d}z_1} \\
&= \int_{\frac{Z(\mathbb{Q}_p) \times G_1(\mathbb{Q}_p)}{Z_1(\mathbb{Q}_p)}} \chi_b^{Z \times G_1} \int_{Z_1(\mathbb{Q}_p)} f_{Z \times G_1, \mu', \alpha}(zz_1, hz_1) \mathrm{d}z_1 (\theta_{\pi_0} \omega_0)(z,h) \frac{\mathrm{d}(z,h)}{\mathrm{d}z_1}. \qquad (5.3)
\end{aligned}$$

We claim that

$$\int_{Z_1(\mathbb{Q}_p)} f_{Z \times G_1, \mu', \alpha}(zz_1, hz_1) \mathrm{d}z_1 = f_{G,\mu,\alpha}(z,h). \qquad (5.4)$$

The map $Z \times G_1 \to G$ is surjective on $\mathbb{Z}_p$-points, and therefore the function

$$\int_{Z_1} f_{Z \times G_1, \mu', \alpha}(zz_1, hz_1) \mathrm{d}z_1$$

is $G(\mathbb{Z}_p)$-spherical. Therefore, to show that Equation (5.4) is true, it suffices to show that the Satake transforms of these functions agree. We compute the Satake transform of the left

hand side:

$$\begin{aligned}\delta_{P_0}^{-1}\int_{N_0(\mathbb{Q}_p)}\int_{Z_1(\mathbb{Q}_p)} & f_{Z\times G_1,\mu',\alpha}(zz_1n_0,hz_1n_0)\mathrm{d}z_1\mathrm{d}n_0 \\ &= \delta_{P_0}^{-1}\int_{Z_1(\mathbb{Q}_p)}\int_{N_0(\mathbb{Q}_p)} f_{Z\times G_1,\mu',\alpha}(zz_1n_0,hz_1n_0)\mathrm{d}n_0\mathrm{d}z_1 \\ &= \int_{Z_1(\mathbb{Q}_p)}\delta_{P_0}^{-1}\int_{N_0(\mathbb{Q}_p)} f_{Z\times G_1,\mu',\alpha}(zz_1n_0,hz_1n_0)\mathrm{d}n_0\mathrm{d}z_1 \\ &= \int_{Z_1(\mathbb{Q}_p)}\left(f_{Z\times G_1,\mu',\alpha}\right)^{(P_0)}(zz_1,hz_1)\mathrm{d}z_1\end{aligned}$$

By Definition 4.1 the last expression is equal to $f_{G,\mu,\alpha}^{(P_0)}(z,h)$. This proves Equation (5.4). We may continue with Equation (5.3) to obtain

$$\int_{\frac{Z(\mathbb{Q}_p)\times G_1(\mathbb{Q}_p)}{Z_1(\mathbb{Q}_p)}}\chi_b^{Z\times G_1}f_{G,\mu,\alpha}\theta_{\pi_0}\omega_0\frac{\mathrm{d}(z,h)}{\mathrm{d}z_1}.$$

Now ω_0 is of finite order by assumption, and the function $f_{G\mu\alpha}$ restricted to $\mathbb{Q}_p{}^\times\cong A(\mathbb{Q}_p)\subset Z(\mathbb{Q}_p)$ is the characteristic function of $p^{-\alpha}\mathbb{Z}_p^\times$. For α sufficiently divisible this is then, up to normalization of Haar measures, just the trace $\mathrm{Tr}(\chi_b^{Z\times G_1}f_{Z\times G_1,\mu',\alpha},\pi_0)$. This proves that $\mathrm{Tr}(\chi_b^G f_{G,\mu,\alpha},\Pi)$ and $\mathrm{Tr}(\chi_b^{Z\times G_1}f_{Z\times G_1,\mu',\alpha},\pi_0)$ differ by a positive, non-zero, scalar. This completes the proof. □

6. Global extension

In this section we prove a technical proposition concerning the restriction of automorphic representations of G to the subgroup $G_1\subset G$ (the kernel of the factor of similitudes). Recall that we have the surjection $G_1\times Z\twoheadrightarrow G$.

Proposition 6.1. *Let Π be a cuspidal automorphic representation of $G(\mathbb{A})$, then its restriction to the group $G_1(\mathbb{A})\times Z(\mathbb{A})$ contains a cuspidal automorphic representation of $G_1(\mathbb{A})\times Z(\mathbb{A})$.*

Remark. The proof we give here is copied from Clozel's article [7, p. 137]; cf. Labesse-Schwermer [24, p. 391].

Proof of Proposition 6.1. Let $A\subset Z$ be the split center. Define $\mathbb{G}_1$ to be the subset $\mathbb{G}_1:=A(\mathbb{A})G(\mathbb{Q})G_1(\mathbb{A})\subset G(\mathbb{A})$. Then $\mathbb{G}_1$ is a subgroup because $G(\mathbb{Q})$ normalizes $G_1(\mathbb{A})$. Furthermore it is easy to see that $\mathbb{G}_1$ is closed in $G(\mathbb{A})$, and we have $[A(\mathbb{A})G_1(\mathbb{A})]\cap G(\mathbb{Q})=A(\mathbb{Q})G_1(\mathbb{Q})\subset G(\mathbb{A})$ (cf. Clozel [Lemme 5.8, *loc. cit.*]). Let χ be the central character of Π; and let ε be the restriction of χ to $G_1(\mathbb{A})\times A(\mathbb{A})$. Let ρ_0 be the representation of $G_1(\mathbb{A})$ on the space $L_0^2(G_1(\mathbb{Q})\backslash G_1(\mathbb{A}),\varepsilon)$ of cuspidal functions that transform under $G_1(\mathbb{A})$ via ε. We extend the representation ρ_0 to a representation of $\mathbb{G}_1$ by defining: $\rho_1(z\gamma x)f(y)=\chi(z)f(\gamma^{-1}y\gamma x)$, for $z\in A(\mathbb{A}),\gamma\in G(\mathbb{Q}),x\in G_1(\mathbb{A}),y\in G_1(\mathbb{A})$. We do not copy the verification that this representation is well-defined [*loc. cit*, 5.16]. Define the representation $\rho=\mathrm{Ind}_{\mathbb{G}_1}^{G(\mathbb{A})}(\rho_1)$ of

$G(\mathbb{A})$. A computation shows that ρ is isomorphic to the representation of $G(\mathbb{A})$ on the space $L^2_0(G(\mathbb{Q})\backslash G(\mathbb{A}), \chi)$ of functions on $G(\mathbb{Q})\backslash G(\mathbb{A})$ which transform via χ under the action of $A(\mathbb{A})$. Consequently, if Π occurs in the representation $\mathrm{Ind}_{\mathbb{G}_1}^{G(\mathbb{A})}(\rho_1)$, then its restriction to $\mathbb{G}_1$ will contain irreducible $\mathbb{G}_1$-subrepresentations of ρ_1. $\square$

7. The isolation argument

Let Sh_K be a Shimura variety of PEL-type of type (A), and let G be the corresponding unitary group of similitudes over $\mathbb{Q}$ (cf. Equation (2.1)). We write E for the reflex field and we let p be a prime of good reduction[3]. Let $b \in B(G_{\mathbb{Q}_p}, \mu)$ be an admissible isocrystal. Let $\mathfrak{p}$ be a prime of the reflex field E above p. Let $\mathbb{F}_q$ be the residue field of E at $\mathfrak{p}$. Let $\mathrm{Sh}^b_{K,\mathfrak{p}}$ be the corresponding Newton stratum of $\mathrm{Sh}_{K,\mathfrak{p}}$, a locally closed subvariety of $\mathrm{Sh}_{K,\wp}$ over $\mathbb{F}_q$ [29].

Let α be a positive integer. We fix an embedding $E_\mathfrak{p} \subset \overline{\mathbb{Q}}_p$ and we write $E_{\mathfrak{p},\alpha}$ for the extension of the field $E_\mathfrak{p}$ of degree α inside $\overline{\mathbb{Q}}_p$.

Theorem 7.1 (Wedhorn-Viehmann)**.** *The variety* $\mathrm{Sh}^b_{K,\mathfrak{p}}$ *is not empty.*

Remark. In the statement of the above theorem we have not been precise about the form of the compact open subgroup $K \subset G(\mathbb{A}_\mathrm{f})$. Note however that for any pair (K, K') of compact open subgroups, which are hyperspecial at p, we have the finite étale morphisms $\mathrm{Sh}_K \leftarrow \mathrm{Sh}_{K\cap K'} \to \mathrm{Sh}_{K'}$ which respect the Newton stratification modulo $\mathfrak{p}$. Therefore, showing that the Newton stratum is non-empty for one K is equivalent to showing that it is true for all K.

Proof. Fix an integer α, sufficiently divisible, so that the conclusion of Proposition 5.1 is true and so that α is even. We start with the formula of Kottwitz. We write ϕ_α for the function $\phi_{G,\mu,\alpha}$ from the previous section[4] on $G(E_{\mathfrak{p},\alpha})$. Similarly $f_\alpha := f_{G,\mu,\alpha}$. We pick a prime $\ell \neq p$ and fix an isomorphism $\overline{\mathbb{Q}}_\ell \cong \mathbb{C}$ (and suppress it from all notations). Let ξ be an irreducible complex (algebraic) representation of G, and write $\mathcal{L}$ for the corresponding ℓ-adic local system on the Shimura tower. Then the Kottwitz formula states:

(7.1)
$$\sum_{x' \in \mathrm{Fix}^b_{f^p\times\Phi^\alpha_\mathfrak{p}}(\overline{\mathbb{F}}_q)} \mathrm{Tr}(f^p \times \Phi^\alpha_\mathfrak{p}, \iota^*(\mathcal{L})_x) = |\,\mathrm{Ker}^1(\mathbb{Q}, G)| \sum_{(\gamma_0;\gamma,\delta)} c(\gamma_0;\gamma,\delta) O_\gamma(f^{\infty p}) \mathrm{TO}_\delta(\phi_\alpha) \,\mathrm{Tr}\, \xi_\mathbb{C}(\gamma_0),$$

where $\mathrm{Fix}^b_{f^p\times\Phi^\alpha_\mathfrak{p}(\overline{\mathbb{F}}_q)}$ is the set of fixed points of the Hecke correspondence $f^p \times \Phi^\alpha_\mathfrak{p}$ acting on $\mathrm{Sh}^b_{K,\mathbb{F}_q}$, and where the sum ranges of the Kottwitz triples $(\gamma_0; \gamma, \delta)$ with the additional condition that the isocrystal defined by δ is equal to b. In Equation (7.1) the map ι is the embedding of $\mathrm{Sh}^b_{K,\mathbb{F}_q}$ into $\mathrm{Sh}_{K,\mathbb{F}_q}$.

[3]Here '*good reduction*' is in the sense of Kottwitz [18, §6]; in particular K decomposes into a product $K = K_p K^p$ with $K_p \subset G(\mathbb{Q}_p)$ hyperspecial.

[4]Where the notation E_α from that section should be replaced with $E_{\mathfrak{p},\alpha}$, and similarly F^+ of that section should be replaced by $F^+ \otimes \mathbb{Q}_p = \prod_\wp F^+_\wp$, where $\wp$ ranges over the places above p.

We may rewrite the right hand side of Equation (7.1) as

$$|\operatorname{Ker}^1(\mathbb{Q},G)| \sum_{(\gamma_0;\gamma,\delta)} c(\gamma_0;\gamma,\delta)\cdot O_\gamma(f^{\infty p})\mathrm{TO}_\delta(\chi_{\sigma b}^{G(E_{\mathfrak{p},\alpha})}\phi_\alpha)\operatorname{Tr}\xi_{\mathbb{C}}(\gamma_0), \tag{7.2}$$

where now the sum ranges over *all* Kottwitz triples and where $\chi_{\sigma b}^{G(E_{\mathfrak{p},\alpha})}$ is the characteristic function on $G(E_{\mathfrak{p},\alpha})$ such for each element $\delta \in G(E_{\mathfrak{p},\alpha})$ we have $\chi_{\sigma b}^{G(E_{\mathfrak{p},\alpha})}(\delta) = 1$ if and only if the conjugacy class $\gamma = \mathcal{N}(\delta)$ satisfies $\Phi(\gamma) = \lambda\overline{\nu}$ for some positive real number $\lambda \in \mathbb{R}_{>0}^\times$. Assume that $(\gamma_0;\gamma,\delta)$ is such that the corresponding term $c(\gamma_0;\gamma,\delta)O_\gamma(f^{\infty p})\mathrm{TO}_\delta(\chi_{\sigma b}^{G(E_{\mathfrak{p},\alpha})}\phi_\alpha)\operatorname{Tr}\xi_{\mathbb{C}}(\gamma_0)$ is non-zero. Then, by the proof of Kottwitz [16], we know that the triple $(\gamma_0;\gamma,\delta)$ arises from some virtual Abelian variety with additional PEL-type structures. In particular the isocrystal defined by δ must lie in the subset $B(G_{\mathbb{Q}_p},\mu) \subset B(G_{\mathbb{Q}_p})$. Thus its end point is determined. We have $\gamma = \mathcal{N}(\delta)$ and $\Phi(\gamma) = \lambda\overline{\nu}_b$ for some λ (Proposition 1.1). Therefore the isocrystal defined by δ must be equal to b. Thus the above sum precisely counts Abelian varieties with additional PEL type structures over $\mathbb{F}_{q^\alpha}$ such that their isocrystal is equal to b.

We have to show that the sum in Equation (7.2) is non-zero. Let $\mathcal{E}$ be the (finite) set of endoscopic groups H associated to G and unramified at all places where the data (G,K) are unramified. By the stabilization argument of Kottwitz [17], we have that the expression in Equation (7.2) is equal to the stable sum

$$\sum_{\mathcal{E}} \iota(G,H)\cdot \mathrm{ST}_e^*((\chi_b^G f_\alpha)^H), \tag{7.3}$$

where $(\chi_b^G f_\alpha)^H$ are the transferred functions, whose existence is guaranteed by the fundamental lemma, the $*$ in ST_e^* means that one only considers stable conjugacy classes satisfying a certain regularity condition (which is empty in case H is a maximal endoscopic group), and finally $\iota(G,H)$ is a constant depending on the endoscopic group (cf. [*loc. cit.*] for the definition).

We will consider only functions such that the transfer $(\chi_b^G f_\alpha)^H$ vanishes for proper endoscopic groups, and therefore we may ignore the regularity condition[5]. Thus, Equation (7.3) simplifies for such functions and gives the equation:

$$\sum_{x' \in \mathrm{Fix}^b_{f^p\times\Phi_{\mathfrak{p}}^\alpha}(\overline{\mathbb{F}}_q)} \operatorname{Tr}(f^p \times \Phi_{\mathfrak{p}}^\alpha, \iota^*(\mathcal{L})_x) = \sum_{\mathcal{E}} \iota(G,H)\mathrm{ST}_e((\chi_b^G f_\alpha)^H). \tag{7.4}$$

Visibly, if we show that the left hand side of Equation (7.4) is non-zero for some Hecke operator f^p, then the variety $\mathrm{Sh}^b_{K,\mathbb{F}_q}$ is non-empty. We will show that the right hand side of Kottwitz's formula does not vanish for some choice of K^p and some choice of f^p.

We write G_0^*, G_1^*, G^* for the quasi-split inner forms of G_0, G_1, and G respectively (we remind the reader that G_0 is defined over F^+ and that $G_1 = \mathrm{Res}_{F^+/\mathbb{Q}}G_0$). The group G^* is the maximal endoscopic group of G. Let $\{x_1, x_2, \ldots, x_d\}$ be the set of prime numbers such

[5] In fact, due to the form of the function f_∞ we have $\mathrm{ST}_e^* = \mathrm{ST}_e$, see [28, thm 6.2.1] or [8, (2.5)].

that the group $G_{\mathbb{Q}_{x_i}}$ is ramified. For v a prime number with $v \notin \{x_1, x_2, \ldots, x_d\}$ the local group $G_{\mathbb{Q}_v}$ is quasi-split, and therefore we may (and do) identify it with the group $G^*_{\mathbb{Q}_v}$. Below we will transfer functions from the group $G(\mathbb{A})$ to the group $G^*(\mathbb{A})$; at the places v with $v \notin \{x_1, x_2, \ldots, x_d, \infty\}$ we have $G(\mathbb{Q}_v) = G^*(\mathbb{Q}_v)$ and using this identification we may (and do) take $(h_v)^{G^*(\mathbb{Q}_v)} = h_v$ for any $h_v \in \mathcal{H}(G(\mathbb{Q}_v))$.

To help the reader understand what we do below at the places x_i (and *why* we do this), let us interrupt this proof with a general remark on the fundamental lemma. It is important to realise that if $v = x_i$ is one of the bad places, then the fundamental lemma does guarantee the *existence* of the transferred function $h_v \rightsquigarrow (h_v)^{G^*(\mathbb{Q}_v)}$; however, in its current state, the fundamental lemma does <u>not</u> give an explicit description of a transferred function $(h_v)^{G^*(\mathbb{Q}_v)}$. The fundamental lemma only gives explicit transfer in case the group is unramified and the level is hyperspecial. For us, the fact that the function $(h_v)^{G^*(\mathbb{Q}_v)}$ is not explicit could introduce signs and cancellations that we cannot control; making it hard to show that expressions such as the one in Equation (7.9) do not vanish. In the argument below we solve the issue by taking h_v to be a pseudocoefficient of the Steinberg representation. For these functions an explicit transfer is known (the transfer is again a pseudocoefficient of the Steinberg representation) and therefore we will be able to control the signs and avoid cancellations. This ends the remark, let us now continue with the proof.

We are going to construct an automorphic representation Π_0 of G^* which has particularly nice properties. From this point onward we take ξ to be a fixed, sufficiently regular complex representation (in the sense of [10, Hyp. (1.2.3)]). We also assume that ξ defines a coefficient system of weight 0 (cf. [8]), and even better that ξ is trivial on the center of G^*. Fix three additional, different, prime numbers p_1, p_2, p_3 ($\neq p$) such that the group $G_{\mathbb{Q}_{p_i}}$ is split for $i = 1, 2, 3$. Let Π_{0,p_1} be a cuspidal representation of the group $G(\mathbb{Q}_{p_1}) = G^*(\mathbb{Q}_{p_1})$. Let $A(\mathbb{R})^+$ be the topological neutral component of the set of real points of the split center A of G. We apply a theorem of Clozel and Shin [5, 32] to find an automorphic representation $\Pi_0 \subset L^2_0(G^*(\mathbb{Q})A(\mathbb{R})^+ \backslash G^*(\mathbb{A}))$ of $G^*(\mathbb{A})$ with:

(1) $\Pi_{0,\infty}$ is in the discrete series and is ξ-cohomological;
(2) $\Pi_{0,p}$ lies in the class $\mathfrak{R}(b)$ (cf. Definition 4.12);
(3) Π_{0,p_1} lies in the *inertial orbit*[6] $\mathcal{I}(\Pi_{0,p_1})$ of Π_{0,p_1} at p_1;
(4) Π_{0,p_2} is isomorphic to the Steinberg representation (up to an unramified twist of finite order);
(5) Π_{0,x_i} is isomorphic to an unramified twist (of finite order) of the Steinberg representation of $G(\mathbb{Q}_{x_i})$ (for $i = 1, 2, \ldots, d$);
(6) $\Pi_{0,v}$ is unramified for all primes $v \notin \{p, p_1, p_2, p_3, x_1, x_2, \ldots, x_d\}$;
(7) The central character of Π_0 has finite order.

[6]For the definition of inertial orbit, see [30, V.2.7].

Because the component at p_1 of Π_0 is cuspidal, we see that Π_0 is a cuspidal automorphic representation. Note that (7) is possible because of the condition on the weight of ξ.

We now choose the group $K \subset G(\mathbb{A}_{\mathrm{f}})$, and we will also choose a compact open group K^* in $G^*(\mathbb{A}_{\mathrm{f}})$. Write $S = \{p, p_1, p_2, p_3\}$. Write $S' = \{p, p_1, p_2, p_3, x_1, x_2, \ldots, x_d\}$ for the union of S with the set of all places where the group G is ramified.

The compact open group $K \subset G(\mathbb{A}_{\mathrm{f}})$ is a (any) group with the following properties:

(1) K is a product $\prod_v K_v \subset G(\mathbb{A}_{\mathrm{f}})$ of compact open groups;
(2) for all $v \notin S'$ the group K_v is hyperspecial;
(3) K_p is hyperspecial;
(4) K_{p_3} is sufficiently small so that Sh_K is smooth and $(\Pi_{0,p_3})^{K_{p_3}} \neq 0$;
(5) K_{x_i} is sufficiently small so that the function f_{x_i} is K_{x_i}-spherical;
(6) for all $v \notin \{x_1, x_2, \ldots, x_d\}$ the space $(\Pi_{0,v})^{K_v}$ is non-zero.

The group $K^* \subset G^*(\mathbb{A}_{\mathrm{f}})$ is a (any) group with the following properties:

(1) K^* is a product $\prod_v K_v^* \subset G^*(\mathbb{A}_{\mathrm{f}})$ of compact open groups;
(2) for any prime $v \notin \{x_1, \ldots, x_d\}$ we have $K_v^* = K_v \subset G(\mathbb{Q}_v) = G^*(\mathbb{Q}_v)$;
(3) for all $i \in \{1, 2, \ldots, d\}$ we have $(\Pi_{0,x_i})^{K_{x_i}} \neq 0$;

We now choose the Hecke function $f \in \mathcal{H}(G(\mathbb{A}_{\mathrm{f}}))$. Consider the function $f^{p\infty} \in \mathcal{H}(G(\mathbb{A}_{\mathrm{f}}))$ of the form

$$f^{p\infty} := f_{p_1} \otimes f_{p_2} \otimes f_{p_3} \otimes f_{x_1} \otimes f_{x_2} \otimes \cdots \otimes f_{x_d} \otimes f^{S'}, \tag{7.5}$$

where

- f_{p_1} is a pseudo-coefficient on $G(\mathbb{Q}_{p_1})$ of the representation Π_{p_1};
- f_{p_2} is a pseudo-coefficient of the Steinberg representation of $G(\mathbb{Q}_{p_2})$;
- $f_{p_3} = \mathbf{1}_{K_{p_3}}$;
- f_{x_i} is (essentially) a pseudo-coefficient of the Steinberg representation of $G(\mathbb{Q}_{x_i})$ for $i = 1, 2, \ldots, d$ (see below for the precise statement and the construction);
- Before we can define the function $f^{S'}$ we have to explain a fact: There are only *finitely many* cuspidal automorphic representations $\Pi \subset L^2_0(G^*(\mathbb{Q})A(\mathbb{R})^+ \backslash G^*(\mathbb{A}))$ of G^* whose component at infinity is equal to Π_∞ and which have invariant vectors under the group K. In particular also the set of their possible outside S'-components $\Pi^{S'}$ is finite. Therefore, we may find a function $f^{S'} \in \mathcal{H}(G^*(\mathbb{A}_{\mathrm{f}}^{S'})) = \mathcal{H}(G(\mathbb{A}_{\mathrm{f}}^{S'}))$ whose trace on $\Pi^{S'}$ is equal to 1 if $\Pi^{S'} \cong \Pi_0^{S'}$ and whose trace equals 0 otherwise for all Π with $\Pi_\infty = \Pi_{0,\infty}$ and $\Pi^K \neq 0$. We fix $f^{S'}$ to be a function which has this property.

We need to comment on the pseudo-coefficients f_{x_i}. In the literature these coefficients are usually only constructed for groups under conditions on the center [9, §3.4], such as the group be semi-simple, or with anisotropic center. We have neither of these conditions. Let $x = x_i$ be one of the bad places and write H for the derived group of G, recall that we write Z for the center of G. We write H^* for the derived group of G^* (then H^* is the quasi-split inner

form of H). Observe that the center Z of G is canonically isomorphic with the center of G^* (and the same is true for the centers of H and H^*). Let k be any smooth function on the group $H(\mathbb{Q}_x)$. Let $O_x \subset Z(\mathbb{Q}_x)$ be the maximal compact open subgroup of the center $Z(\mathbb{Q}_x)$ of $G(\mathbb{Q}_x)$. We now define a function $\widetilde{k}$ on the group $G(\mathbb{Q}_x)$. By definition the function $\widetilde{k}$ has support in the open subgroup $Z(\mathbb{Q}_x)H(\mathbb{Q}_x) \subset G(\mathbb{Q}_x)$, and inside this open subgroup the function is given by the formula

$$\widetilde{k}(g) \stackrel{\mathrm{def}}{=} \frac{1}{\#[O_x \cap H(\mathbb{Q}_x)]} \sum_{z \in O_x \cap H(\mathbb{Q}_x)} k(gz). \tag{7.6}$$

Let H^* be the quasi-split inner form of H; then H^* is also the derived group of G^*. By the fundamental lemma we may transfer smooth functions on the group $G(\mathbb{Q}_x)$ to functions on the group $G^*(\mathbb{Q}_x)$, and similarly functions from the group $H(\mathbb{Q}_x)$ to functions on the group $H^*(\mathbb{Q}_x)$. Note that the formula in Equation (7.6) makes sense if we replace H by its quasi-split inner form; thus we also have a construction $k \mapsto \widetilde{k}$ for smooth functions on $H^*(\mathbb{Q}_x)$. The construction in Equation (7.6) is compatible with transfer of functions, i.e. the function $(\widetilde{k})^{G^*(\mathbb{Q}_x)}$ on $G^*(\mathbb{Q}_x)$ has the same stable orbital integrals as the function $\widetilde{\left(k^{H^*(\mathbb{Q}_x)}\right)}$ for all $k \in \mathcal{H}(H(\mathbb{Q}_x))$.

We now take the function k on $H(\mathbb{Q}_x)$ to be a pseudocoefficient of the Steinberg representation, which exists because the center of H is anisotropic. Define $f_x := \widetilde{k}$. Recall that, in case the group has anisotropic center, the transfer of a pseudocoefficient of the Steinberg representation is again a pseudocoefficient of the Steinberg representation. The conclusion is that we may (and do) take the transferred function $(f_x)^{G^*(\mathbb{Q}_x)}$ to be the one obtained from a pseudocoefficient via the construction in Equation (7.6).

We will show that the function $(f_x)^{G^*(\mathbb{Q}_x)}$ is (essentially) a pseudocoefficient of the Steinberg representation. Let us first make this statement precise. Let χ be a character of the group $G(\mathbb{Q}_x)$. The character χ induces a character $\overline{\chi}$ of the cocenter $C(\mathbb{Q}_x)$ of the group $G(\mathbb{Q}_x)$. We call the character χ *unramified* if $\overline{\chi}$ is trivial on the maximal compact open subgroup K_C of $C(\mathbb{Q}_x)$. We claim that $(f_x)^{G^*(\mathbb{Q}_x)}$ has the following two properties:

$$\begin{array}{ll} \forall\, \chi \text{ unramified character}: & \mathrm{Tr}(f^{G^*(\mathbb{Q}_x)}, \mathrm{St}_{G^*(\mathbb{Q}_x)}(\chi)) \neq 0 \\ \forall\, \Pi_x \text{ smooth irreducible}: & \mathrm{Tr}(f^{G^*(\mathbb{Q}_x)}, \Pi_x) \in \mathbb{R}_{\geq 0}. \end{array} \tag{7.7}$$

We first verify second property in Equation (7.7). Let Π_x be a smooth irreducible representation of the group $G^*(\mathbb{Q}_x)$, let θ_{Π_x} be its character. Let $\pi_1, \ldots, \pi_d$ be the irreducible $H^*(\mathbb{Q}_x)Z(\mathbb{Q}_x)$-subrepresentations of Π_x, and let $\theta_1, \ldots, \theta_d$ be their characters. We have $\theta_{\Pi_x} = \sum_{i=1}^d \theta_i$. Then

(7.8)
$$\mathrm{Tr}(f^{G^*(\mathbb{Q}_x)}, \Pi_x) = \int_{H^*(\mathbb{Q}_x)Z(\mathbb{Q}_x)} f^{G^*(\mathbb{Q}_x)}(g) \sum_{i=1}^d \theta_i(g) \mathrm{d}g = \sum_{i=1}^d \int_{H^*(\mathbb{Q}_x)Z(\mathbb{Q}_x)} f^{G^*(\mathbb{Q}_x)}(g)\theta_i(g)\mathrm{d}g.$$

Now use the definition of the function $f^{G^*(\mathbb{Q}_x)}$ from Equation (7.6) to see that the right hand side is just, up to some positive constant, the trace of the pseudocoefficient of the Steinberg representation on the group $H(\mathbb{Q}_x)$ against π_i. In particular this trace is non-negative, and the second statement in Equation (7.7) is true.

We now verify the first statement in Equation (7.7). By construction the function $f^{G^*(\mathbb{Q}_x)}$ is supported on the inverse image of K_C in G. Because χ is unramified it is constant on the support of $f^{G^*(\mathbb{Q}_x)}$. Therefore we have $\mathrm{Tr}(f_x^{G^*(\mathbb{Q}_x)}, \mathrm{St}_{G^*(\mathbb{Q}_x)}(\chi)) = \mathrm{Tr}(f_x^{G^*(\mathbb{Q}_x)}, \mathrm{St}_{G^*(\mathbb{Q}_x)})$. We verify that $\mathrm{Tr}(f^{G^*(\mathbb{Q}_x)}, \mathrm{St}_{G^*(\mathbb{Q}_x)}) \neq 0$. Let $P_{0,x}$ be a Borel subgroup of $G^*_{\mathbb{Q}_x}$ and let $P'_{0,x}$ be the pull back of $P_{0,x}$ to $H^*_{\mathbb{Q}_x}$. Let I be the space of locally constant complex valued functions on $G^*(\mathbb{Q}_x)/P_{0,x}(\mathbb{Q}_x)$ and I' be the same space, but then for the group $H^*(\mathbb{Q}_x)$. We can extend any function $h \in I'$ by 0 and this gives us the composition of maps $I' \rightarrowtail I \twoheadrightarrow \mathrm{St}_{G^*(\mathbb{Q}_x)}$. This composition is trivial on the subspaces $C^\infty(H^*(\mathbb{Q}_x)/P(\mathbb{Q}_x)) \subset I'$ for any proper parabolic subgroup P of H^* containing $P^*_{0,x}$. We obtain an $H^*(\mathbb{Q}_x)$-injection $\mathrm{St}_{H^*(\mathbb{Q}_x)} \rightarrowtail \mathrm{St}_{G^*(\mathbb{Q}_x)}$. It then follows from Equation (7.8) that $\mathrm{Tr}(f^{G^*(\mathbb{Q}_x)}, \mathrm{St}_{G^*(\mathbb{Q}_x)}) \neq 0$.

We have now completed the definition of the components f_{x_i}, thus also the definition of the Hecke operator $f^{p\infty}$ is complete (see Equation (7.5)). We emphasize that at the primes $v \notin \{x_1, x_2, \ldots, x_d\}$ we take $(f_v)^{G^*(\mathbb{Q}_v)} = f_v$ (we have $G^*(\mathbb{Q}_v) = G(\mathbb{Q}_v)$) and at the primes $v \in \{x_1, x_2, \ldots, x_d\}$ we control the traces of the transferred function $(f_v)^{G^*(\mathbb{Q}_v)}$ against smooth representations via the conclusion in Equation (7.7).

Due to the cuspidal component f_{p_1} of f^p, the trace formula simplifies. Because f_{p_2} is stabilizing (Labesse [22]), the contribution of the proper endoscopic groups are zero, and the right hand side of Equation (7.4) becomes a sum of the form

$$\sum_{\Pi} m(\Pi)\,\mathrm{Tr}((f_\infty f^p)^{G^*(\mathbb{A}^p)}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha), \Pi), \tag{7.9}$$

where Π ranges over cuspidal automorphic representations of $G^*(\mathbb{A})$, and $m(\Pi)$ is the multiplicity of Π in the discrete spectrum of $G^*(\mathbb{A})$ with trivial central character on $A(\mathbb{R})^+$ (recall that A is both the split center of the group G as well as the group G^*). Here we are applying the simple trace formula of Arthur [2, Cor. 23.6] (cf. proof of [1, thm 7.1]), and the fact that, due to the presence of the pseudocoefficients in the Hecke operator, the correcting term in Arthur's formula vanishes. The sum in Equation (7.9) expands to the sum

$$\sum m(\Pi)\,\mathrm{Tr}(f_\infty^{G^*(\mathbb{R})}, \Pi_\infty)\,\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}, \Pi_p)\dim\left((\Pi_{p_3})^{K_{p_3}}\right)\prod_{i=1}^{d}\mathrm{Tr}(f_{x_i}^{G^*(\mathbb{Q}_{x_i})}, \Pi_{x_i}), \tag{7.10}$$

where Π ranges over the irreducible subspaces of $L^2_0(A(\mathbb{R})^+G^*(\mathbb{Q})\backslash G^*(\mathbb{A})/K)$ such that

- $\Pi^{S'} \cong \Pi_0^{S'}$;
- Π_{p_1} lies in the inertial orbit $\mathcal{I}(\Pi_{p_1})$ of the representation Π_{p_1};
- Π_{p_2} is, up to unramified twist, isomorphic to the Steinberg representation of $G(\mathbb{Q}_{p_2})$;
- Π_{x_i} is such that $\mathrm{Tr}(f_{x_i}, \Pi_{x_i}) \neq 0$.

By Proposition 6.1 we may find a cuspidal automorphic representation π_0 of $G_1^*(\mathbb{A})$ contained in Π_0. Let now Π be an automorphic representation of $G^*(\mathbb{A})$ which contributes to Equation (7.10). Thus we have $\Pi^{S'} \cong \Pi_0^{S'}$. Let π be a cuspidal automorphic subrepresentation of $\mathrm{Res}_{[G_1^* \times Z](\mathbb{A})}(\Pi)$ (Proposition 6.1). Then, possibly after enlarging S' to a larger finite set S'', we have that π and Π are unramified for all places outside the set S''. At the unramified places $v \notin S''$ the representation $\mathrm{Res}_{[G_1^* \times Z](\mathbb{Q}_v)}(\Pi_{0,v})$ contains exactly one unramified representation: $\pi_{0,v}$. Therefore we have $(\pi)^{S''} \cong (\pi_0)^{S''}$.

We now apply base change. Observe that π has the following properties:

(1) π is cuspidal;
(2) π_∞ is in the discrete series;
(3) π_{p_1} is cuspidal;
(4) π_{p_2} is an unramified twist of the Steinberg representation.

Consider the group $G_0^{*+} := \mathrm{Res}_{F/F^+}(G_{0,F}^*)$. Let $\mathbb{A}_{F^+} := \mathbb{A} \otimes_{\mathbb{Q}} F^+$ and $\mathbb{A}_F := \mathbb{A} \otimes_{\mathbb{Q}} F$. Then $G_0^{*+}(\mathbb{A}_{F^+}) = G_0^*(\mathbb{A}_F)$. Because of the above properties (1), ..., (4), we may base change π to an automorphic representation $BC(\pi)$ of $G_0^{*+}(\mathbb{A}_{F^+})$. Here we are using Corollary 5.3 from Labesse [23] to see that π has a weak base change, and then the improvement of the statement at Theorem 5.9 of [*loc. cit.*], which states that[7], at the places where the unitary group is quasi-split (so in particular at p) the (local) base change of the representation π_p is the representation $BC(\pi)_p$. By the same argument the base change $BC(\pi_0)$ exists as well. Note that by strong multiplicity one for the group G_0^{*+} we have $BC(\pi_\wp) \cong BC(\pi_{0,\wp})$ for all F^+-places $\wp$ above p.

We give the final argument when F/F^+ is inert at the F^+-place $\wp | p$, the case of the general linear groups being easier.

The representation $\pi_\wp$ is of the form $\mathrm{Ind}_{P(F_\wp^+)}^{G_1(F_\wp^+)}(\rho_\wp)$ because π_p lies in the set $\mathfrak{R}_1(b)$. In this induction the parabolic subgroup P has Levi component M with $M(\mathbb{Q}_p) = M_{\wp,1} \times M_{\wp,2}$ with $M_{\wp,1}$ a general linear group and $M_{\wp,2}$ is a unitary group. The representation $\rho_\wp$ decomposes into $\rho_\wp \cong \rho_{\wp,1} \otimes \rho_{\wp,2}$, where $\rho_{\wp,1}$ is a generic unramified representation of $M_{\wp,1}$ and $\rho_{\wp,2}$ is an unramified twist of the Steinberg representation of $M_{\wp,2}$. The base change is compatible with parabolic induction, the base change of a generic unramified representation is again unramified [26], the base change of a twist of the Steinberg representation is again a twist of the Steinberg representation [27]. Thus the representation $BC(\pi_\wp) \cong BC(\pi_{0,\wp})$ is an induction from a representation of the form

$$\left(\chi_1, \chi_2, \ldots, \chi_{a_\wp}, \mathrm{St}_{\mathrm{GL}_b(F_\wp^+)}, \overline{\chi}_{a_\wp}^{-1}, \overline{\chi}_{a_\wp - 1}^{-1}, \ldots, \overline{\chi}_1^{-1}\right)$$

[7]Here we should remark that Labesse works under the assumption that $[F^+ : \mathbb{Q}]$ is of degree at least 2, and that we do not have this assumption. However, Labesse only needs his assumption to apply the simple trace formula. For our representation π the assumption of Labesse is not needed, because we have an auxiliary place ($v = p_1$) where the representation π is cuspidal.

where $a_\wp = \mathrm{Rank}(M_{\wp,1})$ and $b_\wp = n - a_\wp$. Consequently, we have the character relations

$$\Theta_{\pi_{0,\wp}} \circ \mathcal{N} = \Theta_{\pi_\wp} \circ \mathcal{N}, \tag{7.11}$$

where $\mathcal{N}$ is the norm mapping from $G_0^{*+}(F_\wp^+)$ to $G_0^*(F_\wp^+)$. The norm mapping $\mathcal{N}$ from θ-conjugacy classes in $G_0^{*+}(F_\wp^+)$ to $G_0^*(F_\wp^+)$ is surjective for the semi-simple conjugacy classes [31, Prop. 3.11(b)]. It follows that the characters $\Theta_{\pi_\wp}$ and $\Theta_{\pi_{0,\wp}}$ coincide on $G_0(F_\wp^+)$. By Proposition 5.1 there is a positive constant $C_\Pi \in \mathbb{R}_{>0}$ such that (for α sufficiently divisible)

$$\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p) = C_\Pi \,\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_{0,p}). \tag{7.12}$$

Remark: To find Equation (7.12) we applied Proposition (5.1) two times: first to compare $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p)$ with $\mathrm{Tr}(\chi_b^{G_1\times Z} f_\alpha^{G_1 \times Z}, \pi_p)$, and then to compare $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_{0,p})$ with $\mathrm{Tr}(\chi_b^{G_1\times Z} f_\alpha^{G\times Z_1}, \pi_{0,p})$.

We may now complete the proof. We return to Equation (7.10):

$$\sum m(\Pi)\,\mathrm{Tr}(f_\infty^{G^*(\mathbb{R})}, \Pi_\infty)\,\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p)\dim\left((\Pi_{p_3})^{K_{p_3}}\right)\prod_{i=1}^d \mathrm{Tr}(f^{G^*(\mathbb{Q}_{x_i})}, \Pi_{x_i}), \tag{7.13}$$

where Π ranges over the irreducible subspaces of $L_0^2(A(\mathbb{R})^+G^*(\mathbb{Q})\backslash G^*(\mathbb{A}))$ satisfying the conditions listed below Equation (7.10). The following 6 facts have been established:

(1) The sum in Equation (7.13) is non-empty because Π_0 occurs in it (by the Propositions 4.8 and 4.10, and the Equation (7.7), the term corresponding to Π_0 in the Sum (7.13) is non-zero).
(2) The multiplicity $m(\Pi)$ is a positive real number.
(3) For any Π in Equation (7.13) with $\Pi \not\cong \Pi_0$ we must have $\mathrm{Tr}(f_\infty^{G^*(\mathbb{R})}, \Pi_\infty) = \mathrm{Tr}(f_\infty^{G^*(\mathbb{R})}, \Pi_{0,\infty})$ (here we are using that ξ is sufficiently regular).
(4) By Equation (7.12) the trace $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p)$ equals $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_{0,p})$ up to the positive number C_Π.
(5) The dimensions $\dim\left((\Pi_{p_3})^{K_{p_3}}\right)$ and $\dim\left((\Pi_{0,p_3})^{K_{p_3}}\right)$ differ by a positive real number.
(6) The product $\prod_{i=1}^d \mathrm{Tr}(f^{G^*(\mathbb{Q}_{x_i})}, \Pi_x)$ is a non-negative real number for all automorphic representations Π which contribute to Equation (7.13) (here we use the conclusion at Equation (7.7)).

(facts (2) and (5) are trivial). From facts (1), (2), …, (6) we conclude that Equation (7.13) must be non-zero. This completes the proof. □

References

[1] J. Arthur. The invariant trace formula. II. Global theory. *J. Amer. Math. Soc.*, 1(3):501–554, 1988.

[2] J. Arthur. An introduction to the trace formula. In *Harmonic analysis, the trace formula, and Shimura varieties*, volume 4 of *Clay Math. Proc.*, pages 1–263. Amer. Math. Soc., Providence, RI, 2005.

[3] J. Arthur and L. Clozel. *Simple algebras, base change, and the advanced theory of the trace formula*, volume 120 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1989.

[4] N. Bourbaki. *Lie groups and Lie algebras. Chapters 4–6*. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 2002. Translated from the 1968 French original by Andrew Pressley.

[5] L. Clozel. On limit multiplicities of discrete series representations in spaces of automorphic forms. *Invent. Math.*, 83(2):265–284, 1986.

[6] L. Clozel. The fundamental lemma for stable base change. *Duke Math. J.*, 61(1):255–302, 1990.

[7] L. Clozel. Représentations galoisiennes associées aux représentations automorphes autoduales de $\mathrm{GL}(n)$. *Inst. Hautes Études Sci. Publ. Math.*, (73):97–145, 1991.

[8] L. Clozel. Purity reigns supreme. *Preprint*, 2011.

[9] L. Clozel and G. Chenevier. Corps des nombres peu ramifies et formes automorphes autoduales. *Journal of the A.M.S.*, 22:467–519, 2009.

[10] L. Clozel, M. Harris, and J.-P. Labesse. Construction of automorphic Galois representations, I. In *On the stabilization of the trace formula*, volume 1 of *Stab. Trace Formula Shimura Var. Arith. Appl.*, pages 497–527. Int. Press, Somerville, MA, 2011.

[11] M. Demazure. *Lectures on p-divisible groups*, volume 302 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. Reprint of the 1972 original.

[12] M. Harris and R. Taylor. *The geometry and cohomology of some simple Shimura varieties*, volume 151 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2001. With an appendix by Vladimir G. Berkovich.

[13] R. E. Kottwitz. Rational conjugacy classes in reductive groups. *Duke Math. J.*, 49(4):785–806, 1982.

[14] R. E. Kottwitz. Shimura varieties and twisted orbital integrals. *Math. Ann.*, 269(3):287–300, 1984.

[15] R. E. Kottwitz. Isocrystals with additional structure. *Compositio Math.*, 56(2):201–220, 1985.

[16] R. E. Kottwitz. Shimura varieties and λ-adic representations. In *Automorphic forms, Shimura varieties, and L-functions, Vol. I (Ann Arbor, MI, 1988)*, volume 10 of *Perspect. Math.*, pages 161–209. Academic Press, Boston, MA, 1990.

[17] R. E. Kottwitz. On the λ-adic representations associated to some simple Shimura varieties. *Invent. Math.*, 108(3):653–665, 1992.

[18] R. E. Kottwitz. Points on some Shimura varieties over finite fields. *J. Amer. Math. Soc.*, 5(2):373–444, 1992.

[19] R. E. Kottwitz. Isocrystals with additional structure. II. *Compositio Math.*, 109(3):255–339, 1997.

[20] A. Kret. The basic stratum of some simple Shimura varieties. *Preprint*, 2011.

[21] A. Kret. The basic stratum of Kottwitz's arithmetical varieties. *Preprint*, 2012.

[22] J.-P. Labesse. Cohomologie, stabilisation et changement de base. *Astérisque*, (257):vi+161, 1999. Appendix A by Laurent Clozel and Labesse, and Appendix B by Lawrence Breen.

[23] J.-P. Labesse. Changement de base CM et séries discrètes. In *On the stabilization of the trace formula*, volume 1 of *Stab. Trace Formula Shimura Var. Arith. Appl.*, pages 429–470. Int. Press, Somerville, MA, 2011.

[24] J.-P. Labesse and J. Schwermer. On liftings and cusp cohomology of arithmetic groups. *Invent. Math.*, 83(2):383–401, 1986.

[25] J.-P. Labesse and J.-L. Waldspurger. La formule des traces tordue d'apres le friday morning seminar. *Preprint*, 2012.

[26] A. Mínguez. Unramified representations of unitary groups. In *On the stabilization of the trace formula*, volume 1 of *Stab. Trace Formula Shimura Var. Arith. Appl.*, pages 389–410. Int. Press, Somerville, MA, 2011.

[27] C. Mœglin. Classification et changement de base pour les séries discrètes des groupes unitaires p-adiques. *Pacific J. Math.*, 233(1):159–204, 2007.

[28] S. Morel. *On the cohomology of certain noncompact Shimura varieties*, volume 173 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2010. With an appendix by Robert Kottwitz.

[29] M. Rapoport and M. Richartz. On the classification and specialization of F-isocrystals with additional structure. *Compositio Math.*, 103(2):153–181, 1996.

[30] D. Renard. *Représentations des groupes réductifs p-adiques*, volume 17 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, 2010.

[31] J. D. Rogawski. *Automorphic representations of unitary groups in three variables*, volume 123 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1990.

[32] S. W. Shin. Automorphic Plancherel density theorem. *to appear in Israel J. Math*, 2011.

[33] J. Tits. Reductive groups over local fields. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1*, Proc. Sympos. Pure Math., XXXIII, pages 29–69. Amer. Math. Soc., Providence, R.I., 1979.

[34] J.-L. Waldspurger. La formule de Plancherel pour les groupes p-adiques (d'après Harish-Chandra). *J. Inst. Math. Jussieu*, 2(2):235–333, 2003.

[35] C.-H. Wang. Representations de Weil pour les groupes de similitudes et changement de base. *thesis, Paris-sud*, 2012.

[36] T. Wedhorn and E. Viehmann. Ekedahl-Oort and Newton strata for Shimura varieties of pel type. *Preprint*, 2011.

Arno Kret
Université Paris-Sud, UMR 8628, Mathématique, Bâtiment 425,
F-91405 Orsay Cedex, France